

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 15:54:19 : Search time 21.7073 Seconds

(without alignments)  
3044.503 Million cell updates/sec

Title: US-09-941-947a-18

Perfect score: 1623

Sequence: 1 MOIVIANPRGFCAGVDRAIE.....GIEEKVVFSLPKELKKHQA 318

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*

2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*

5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*

7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*

9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*

10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*

11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*

12: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*

13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*

14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*

15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*

16: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*

17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*

18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1623	100.0	318	US-09-934-903-24	Sequence 24, Appl
2	1623	100.0	318	US-09-941-847a-18	Sequence 18, Appl
3	1028.5	63.4	349	US-09-975-719-429	Sequence 429, Appl
4	787	48.5	338	US-10-156-761-10747	Sequence 10747, A
5	755	46.5	325	US-09-738-626-4643	Sequence 4643, Ap
6	743.5	45.8	329	US-09-712-363-271	Sequence 271, App
7	733	45.2	335	US-09-712-363-188	Sequence 188, App
8	663	40.9	310	US-10-289-762-1090	Sequence 1090, Ap
9	663	40.9	310	US-10-022-832-36	Sequence 36, Appl
10	598.5	36.9	539	US-10-275-260-5	Sequence 5, Appl
11	111	6.8	917	US-10-156-761-10047	Sequence 10047, A
12	109	6.7	578	US-09-712-363-282	Sequence 282, App
13	109	6.7	925	US-10-369-493-5415	Sequence 5415, App
14	104.5	6.4	501	US-10-369-493-1052	Sequence 1052, Ap
15	104.5	6.4	1130	US-10-369-493-19263	Sequence 19263, A

16	104	6.4	869	12	US-10-369-493-19983	Sequence 19983, A
17	103	6.3	433	9	US-09-815-242-10607	Sequence 10607, A
18	103	6.3	553	12	US-10-033-585-7338	Sequence 7338, Ap
19	102	6.3	1382	10	US-09-728-653-2	Sequence 2, Appl
20	101	6.2	381	12	US-10-369-493-20969	Sequence 20969, A
21	101	6.2	520	12	US-10-369-493-2893	Sequence 2893, Ap
22	100.5	6.2	538	12	US-10-369-493-2950	Sequence 2950, Ap
23	99.5	6.1	1765	15	US-10-037-182-8	Sequence 8, Appl
24	99.5	6.1	1786	9	US-09-873-676-113	Sequence 113, App
25	99.5	6.1	1786	10	US-09-938-275-6	Sequence 6, Appl
26	99.5	6.1	1786	15	US-10-037-182-6	Sequence 6, Appl
27	97.5	6.0	607	12	US-10-368-493-19473	Sequence 19473, A
28	97.5	6.0	921	10	US-09-738-626-5659	Sequence 5659, Ap
29	97	6.0	299	12	US-10-369-493-10700	Sequence 10700, A
30	96.5	5.9	321	9	US-09-815-242-13454	Sequence 13454, A
31	96.5	5.9	1881	12	US-10-033-585-7646	Sequence 7646, Ap
32	96	5.9	952	12	US-10-108-260A-3284	Sequence 3284, Ap
33	95.5	5.9	330	12	US-10-369-493-23413	Sequence 23413, A
34	95.5	5.9	977	12	US-10-342-136-2	Sequence 2, Appl
35	94.5	5.8	759	12	US-10-369-493-18629	Sequence 18629, A
36	94	5.8	496	15	US-10-263-677-9	Sequence 7, Appl
37	94	5.8	496	15	US-10-263-677-9	Sequence 8, Appl
38	94	5.8	547	12	US-09-847-6378-8	Sequence 7917, Ap
39	93.5	5.8	547	12	US-10-369-493-7917	Sequence 135, App
40	93	5.7	563	12	US-10-341-434-135	Sequence 1, Appl
41	93	5.7	573	10	US-09-828-574-1	Sequence 1, Appl
42	93	5.7	573	12	US-10-307-326-1	Sequence 1, Appl
43	93	5.7	573	12	US-10-116-275-141	Sequence 141, App
44	93	5.7	573	12	US-10-367-093-12	Sequence 12, Appl
45	93	5.7	573	15	US-10-046-649-1	Sequence 1, Appl

## ALIGNMENTS

US-09-934-903-24	Sequence 24, Application US/099344903
1	Patent No. US20020102690A1
GENERAL INFORMATION:	
APPLICANT:	Kofias, Mathias
APPLICANT:	Odom, J. Martin
APPLICANT:	Schenzle, Andreas J.
APPLICANT:	No. US20020102690A1ton, Kelley C.
APPLICANT:	Tomb, Jean-Francois
APPLICANT:	Rouviere, Pierre
APPLICANT:	Picateggio, Stephen
APPLICANT:	Cheng, Qiong
TITLE OF INVENTION:	Genes Involved in Isoprenoid Compounds Production
FILE REFERENCE:	CI:646 US NA
CURRENT APPLICATION NUMBER:	US/09/934,903
CURRENT FILING DATE:	2001-08-22
PRIOR APPLICATION NUMBER:	60/229,907
PRIOR FILING DATE:	September 1, 2001
NUMBER OF SEQ ID NOS:	24
SOFTWARE:	Microsoft Office 97
SEQ ID NO 24	
LENGTH:	318
TYPE:	prt
ORGANISM:	Methylobionas 16a
US-09-934-903-24	
Query Match	100.0%; Score 1623; DB 10; Length 318;
Best Local Similarity	100.0%; Pred. No. 2.5e-150;
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
DB	1 MOIVIANPRGFCAGVDRAIEIVDOAIEAFGAPIVREYVHNRTVVGILQKGVFIEEL 60
QY	SDVPVGYSLIFSAGVKEVQOAEERQLTVFDATCELVTKVHQVAKAKGGEVLLIG 120
DB	61 SDVPVGYSLIFSAGVKEVQOAEERQLTVFDATCELVTKVHQVAKAKGGEVLLIG 120

QY 121 HAGHEVEGTMGOYEKCTEGGCIYLVETPEDVNRLLKNNPNDLAVYVOTLLSTMTDKVWV 180  
 DB 121 HAGHEVEGTMGOYEKCTEGGCIYLVETPEDVNRLLKNNPNDLAVYVOTLLSTMTDKVWV 180  
 QY 181 DALREOPPSIKEOKKDDICATONRQDAVHDLAKISDLILVSGSPNSNSNRLREIAVOL 240  
 DB 181 DALREOPPSIKEOKKDDICATONRQDAVHDLAKISDLILVSGSPNSNSNRLREIAVOL 240  
 QY 241 GKPAYLIDTYODLKQDMLEGIENVGVTAAGASAPVLYVQETIDOLKAMGGETTSVRENSGI 300  
 DB 241 GKPAYLIDTYODLKQDMLEGIENVGVTAAGASAPVLYVQETIDOLKAMGGETTSVRENSGI 300  
 QY 301 EEKXVFSIPKELKXKMOA 318  
 DB 301 EEKXVFSIPKELKXKMOA 318

RESULT 2  
 US-09-941-947a-18  
 ; Sequence 18, Application US/09941947A  
 ; Publication No. US20030003528A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brzostowicz, Patricia C.  
 ; APPLICANT: Cheng, Qiong  
 ; APPLICANT: DiCosimo, Deana J.  
 ; APPLICANT: Koffas, Mattheos  
 ; APPLICANT: Miller, Edward S. Jr.  
 ; APPLICANT: Odom, J. Martin  
 ; APPLICANT: Picataggio, Steve  
 ; APPLICANT: Roviello, Pierre B.  
 ; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE  
 ; FILE REFERENCE: CIL903 US NA  
 ; CURRENT APPLICATION NUMBER: US/09/941.947A  
 ; PRIOR FILING DATE: 2001-09-01  
 ; PRIOR APPLICATION NUMBER: 60/229,907  
 ; PRIOR FILING DATE: 2000-09-01  
 ; PRIOR APPLICATION NUMBER: 60/229,858  
 ; NUMBER OF SEQ ID NOS: 60  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 18  
 ; LENGTH: 318  
 ; TYPE: PRT  
 ; ORGANISM: Methylobionas 16a  
 US-09-941-947a-18

Query Match 100.0%; Score 1623; DB 11; Length 318;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-150;  
 Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOVLANRGRGACAGVDAIEIVDOAIEAFGAPIVRRHVNRVVDNLKQKAVFIEEL 60  
 DB 1 MOVLANRGRGACAGVDAIEIVDOAIEAFGAPIVRRHVNRVVDNLKQKAVFIEEL 60  
 QY 61 SDVPVGYLIFSAGHVSKEVQOABEROLTVFDPATCELVTKVHQAQAKOGREVLIG 120  
 DB 61 SDVPVGYLIFSAGHVSKEVQOABEROLTVFDPATCELVTKVHQAQAKOGREVLIG 120  
 QY 121 HAGHEVEGTMGOYEKCTEGGCIYLVETPEDVNRLLKNNPNDLAVYVOTLLSTMTDKVWV 180  
 DB 121 HAGHEVEGTMGOYEKCTEGGCIYLVETPEDVNRLLKNNPNDLAVYVOTLLSTMTDKVWV 180  
 QY 181 DALREOPPSIKEOKKDDICATONRQDAVHDLAKISDLILVSGSPNSNSNRLREIAVOL 240  
 DB 181 DALREOPPSIKEOKKDDICATONRQDAVHDLAKISDLILVSGSPNSNSNRLREIAVOL 240  
 QY 241 GKPAYLIDTYODLKQDMLEGIENVGVTAAGASAPVLYVQETIDOLKAMGGETTSVRENSGI 300  
 DB 241 GKPAYLIDTYODLKQDMLEGIENVGVTAAGASAPVLYVQETIDOLKAMGGETTSVRENSGI 300  
 QY 301 EEKXVFSIPKELKXKMOA 318  
 DB 301 EEKXVFSIPKELKXKMOA 318

DB 301 EEKXVFSIPKELKXKMOA 318

RESULT 3  
 US-09-975-719-429  
 ; Sequence 429, Application US/09975719  
 ; Publication No. US20030022349A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ausubel, Frederick M.  
 ; APPLICANT: Rahme, Laurence G.  
 ; TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
 ; FILE REFERENCE: 00786/361003  
 ; CURRENT APPLICATION NUMBER: US/09/975.719  
 ; PRIOR FILING DATE: 2001-10-10  
 ; PRIOR APPLICATION NUMBER: US 09/199,637  
 ; PRIOR FILING DATE: 1998-11-25  
 ; PRIOR APPLICATION NUMBER: US 60/066,517  
 ; PRIOR FILING DATE: 1997-11-25  
 ; NUMBER OF SEQ ID NOS: 437  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 429  
 ; LENGTH: 349  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-975-719-429

Query Match 63.4%; Score 1028.5; DB 11; Length 349;  
 Best Local Similarity 63.3%; Pred. No. 4.7e-92;  
 Matches 198; Conservative 49; Mismatches 63; Indels 3; Gaps 2;

QY 1 MOVLANRGRGACAGVDAIEIVDOAIEAFGAPIVRRHVNRVVDNLKQKAVFIEEL 60  
 DB 36 MOVLANRGRGACAGVDAIEIVDOAIEAFGAPIVRRHVNRVVDNLKQKAVFIEEL 95  
 QY 61 SDVPVGYLIFSAGHVSKEVQOABEROLTVFDPATCELVTKVHQAQAKOGREVLIG 120  
 DB 96 DQVDNVIIVIFSAGHVSKEVQOABEROLTVFDPATCELVTKVHQAQAKOGREVLIG 155  
 QY 121 HAGHEVEGTMGOYEKCTEGGCIYLVETPEDVNRLLKNNPNDLAVYVOTLLSTMTDKVWV 180  
 DB 156 HAGHEVEGTMGOYEKCTEGGCIYLVETPEDVNRLLKNNPNDLAVYVOTLLSTMTDKVWV 214  
 QY 181 DALREOPPSIKEOKKDDICATONRQDAVHDLAKISDLILVSGSPNSNSNRLREIAVOL 240  
 DB 215 DALREOPPSIKEOKKDDICATONRQDAVHDLAKISDLILVSGSPNSNSNRLREIAVOL 274  
 QY 241 GKPAYLIDTYODLKQDMLEGIENVGVTAAGASAPVLYVQETIDOLKAMGGETTSVRENSGI 300  
 DB 275 GKPAYLIDTYODLKQDMLEGIENVGVTAAGASAPVLYVQETIDOLKAMGGETTSVRENSGI 332  
 QY 301 EEKXVFSIPKELK 313  
 DB 333 BENITFSMPKELR 345  
 RESULT 4  
 US-10-156-761-10747  
 ; Sequence 10747, Application US/10156761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIKAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHIYUKI  
 ; APPLICANT: HATTORI, MASAHIRA  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 248-262  
 ; CURRENT APPLICATION NUMBER: US/10/156.761  
 ; PRIOR FILING DATE: 2002-05-29  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: JP 2001-272697  
PRIOR FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 15109  
SEQ ID NO 10747  
LENGTH: 338  
TYPE: PRT  
ORGANISM: Streptomyces avermitilis  
US-10-156-761-10747

Query Match 48.5%; Score 787; DB 15; Length 338;  
Best Local Similarity 48.1%; Pred. No. 2e-68;  
Matches 153; Conservative 71; Mismatches 86; Indels 8; Gaps 3;

QY 2 QIVANPRGFCAGVRAIEIVDOAIEAFGAPIVYRHEVNNRTVVDGLKQKGAFFIELS 61  
DB 11 RVLLAARGYCAQVRAIVAVKRLBQYCAPIVYRHEVNNRTVVDGLKQKGAFFIELS 70  
QY 62 DVPVGSYLIFSAHGVSAKVEQAEERQLTVPDTCPLVTYVMQVAKAKQREVIILIGH 121  
DB 71 EEPGATVMSFSAHGVSAKVEQAEERQLTVPDTCPLVTYVMQVAKAKQREVIILIGH 130  
QY 122 AGHEVEGSTMQYKCTEGGGIYVETPEDEVNKLKVNPNDAVYVQTTLSTMTDKVMD 181  
DB 131 EGHEVEGTMQYKCTEGGGIYVETPEDEVNKLKVNPNDAVYVQTTLSTMTDKVMD 185  
QY 182 ALRQFPISIKQKDDICVATONRODAVHDLAKISDILIVGSPNSNSNRLREIAVQ 241  
DB 186 ALRQFPISIKQKDDICVATONRODAVHDLAKISDILIVGSPNSNSNRLREIAVQ 245  
QY 242 -KRAYLDITVODLQKMDLEGIEVGVYAGASAPVYVQVYIDOLKXMGGETTSVRENSG 300  
DB 246 ARDYLVDVDEIDEAMLEGVSTVTSASVPEILVEQVLEWLSQRFEDVEIVK-AA 303  
QY 301 EEKVPFSIPKELKQMOA 318  
DB 304 EESITFSLPKELRDLRA 321

RESULT 5  
US-09-738-626-4643  
Sequence 4643, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:  
APPLICANT: NAKAGATA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAOKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738, 626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/151162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 4643  
LENGTH: 325  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4643

Query Match 46.5%; Score 755; DB 10; Length 325;

Best Local Similarity 48.3%; Pred. No. 2.5e-65;  
Matches 152; Conservative 67; Mismatches 86; Indels 10; Gaps 4;

QY 2 QIVANPRGFCAGVRAIEIVDOAIEAFGAPIVYRHEVNNRTVVDGLKQKGAFFIELS 61  
DB 15 RVLLAARGYCAQVRAIVAVKRLBQYCAPIVYRHEVNNRTVVDGLKQKGAFFIELS 74  
QY 62 DVPVGSYLIFSAHGVSAKVEQAEERQLTVPDTCPLVTYVMQVAKAKQREVIILIGH 121  
DB 75 EEPGATVMSFSAHGVSAKVEQAEERQLTVPDTCPLVTYVMQVAKAKQREVIILIGH 134  
QY 122 AGHEVEGSTMQYKCTEGGGIYVETPEDEVNKLKVNPNDAVYVQTTLSTMTDKVMD 179  
DB 135 EGHEVEGTMQYKCTEGGGIYVETPEDEVNKLKVNPNDAVYVQTTLSTMTDKVMD 189  
QY 180 ALRQFPISIKQKDDICVATONRODAVHDLAKISDILIVGSPNSNSNRLREIAVQ 239  
DB 190 ALRQFPISIKQKDDICVATONRODAVHDLAKISDILIVGSPNSNSNRLREIAVQ 249  
QY 240 -KRAYLDITVODLQKMDLEGIEVGVYAGASAPVYVQVYIDOLKXMGGETTSVRENS 298  
DB 250 ARDYLVDVDEIDEAMLEGVSTVTSASVPEILVEQVLEWLSQRFEDVEIVK-AA 307  
QY 299 EEKVPFSIPKELK 313  
DB 308 EESITFSLPKELRDLRA 322

RESULT 6  
US-09-712-363-271  
Sequence 271, Application US/09712363  
Patent No. US20020164588A1  
GENERAL INFORMATION:  
APPLICANT: Eisenberg, David  
APPLICANT: Rotstein, Sergio H.  
APPLICANT: Marcotte, Edward M.  
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS  
FILE REFERENCE: 07419-032001  
CURRENT APPLICATION NUMBER: US/09/712, 363  
CURRENT FILING DATE: 2000-11-13  
PRIOR APPLICATION NUMBER: PCT/US00/02246  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 60/179, 531  
PRIOR FILING DATE: 2000-02-01  
PRIOR APPLICATION NUMBER: 60/117, 844  
PRIOR FILING DATE: 1999-01-29  
PRIOR APPLICATION NUMBER: 60/118, 206  
PRIOR FILING DATE: 1999-02-01  
PRIOR APPLICATION NUMBER: 60/126, 593  
PRIOR FILING DATE: 1999-03-26  
PRIOR APPLICATION NUMBER: 60/134, 093  
PRIOR FILING DATE: 1999-05-14  
PRIOR APPLICATION NUMBER: 60/134, 092  
PRIOR FILING DATE: 1999-05-14  
PRIOR APPLICATION NUMBER: 60/165, 124  
PRIOR FILING DATE: 1999-11-12  
PRIOR APPLICATION NUMBER: 60/165, 086  
PRIOR FILING DATE: 1999-11-12  
NUMBER OF SEQ ID NOS: 292  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 271  
LENGTH: 329  
TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis  
US-09-712-363-271

Query Match 45.8%; Score 743.5; DB 10; Length 329;  
Best Local Similarity 47.6%; Pred. No. 3.4e-64;  
Matches 150; Conservative 60; Mismatches 94; Indels 11; Gaps 3;  
QY 3 QIVANPRGFCAGVRAIEIVDOAIEAFGAPIVYRHEVNNRTVVDGLKQKGAFFIELS 62  
DB 15 RVLLAARGYCAQVRAIVAVKRLBQYCAPIVYRHEVNNRTVVDGLKQKGAFFIELS 74

Db 20 VLLASRSPGAGVERALETYKRLVDVABGVYVRKQIVANTVVAELRDGAVFVEDDE 79  
 Qy 63 V---PVGSYLIFSANGSKEVQOAEEROLYFADATCPULVTKEQVAKAKGREFVL 118  
 Db 80 IPDPBPFGAVVPSAIGVSPAVRAGADERGLQVADATCPULVAKVHEAARFPAARBDTWVF 139  
 Qy 119 IGHAGHEVGTMGQVEKCTEGGGIYLVETPEDEVRLKNNPNDLAVYVQTLSMTDTKV 178  
 Db 140 IGHAGHEVGTMGQVEKCTEGGGIYLVETPEDEVRLKNNPNDLAVYVQTLSMTDTKV 194  
 Qy 179 MDALRBPSPSIXEOKKDDICVATONRQDAVHDLAKISDLIVGSPNSNSNRLRELVAV 238  
 Db 195 VIDLRARFPFLGQPPEDICVATNRQALQSMWECVVLVIGSCNSNSRRLRELVAV 254  
 Qy 239 QLGKPAALIDYQDKODMLEGIEVVGVTAGASAPVLYOEVITDOLKAWGETTSVRENS 298  
 Db 255 RSGTPAYLIDGPPDIPEWELSSVTIGVTAGASAPVLYOEVITDOLKAWGETTSVRENS 312  
 Qy 299 GIEKVFSPKPELK 313  
 Db 313 IATETVRFGLPKQVR 327

RESULT 7  
 US-09-712-363-188  
 ; Sequence 188, Application US/09712363  
 ; Patent No. US20020164588A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eisenberg, David  
 ; APPLICANT: Rotstein, Sergio H.  
 ; APPLICANT: Marcotte, Edward M.  
 ; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
 ; FILE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS  
 ; FILE REFERENCE: 07419-032001  
 ; CURRENT APPLICATION NUMBER: US/09/712,363  
 ; CURRENT FILING DATE: 2000-11-13  
 ; PRIOR APPLICATION NUMBER: PCT/US00/02246  
 ; PRIOR FILING DATE: 2000-01-28  
 ; PRIOR APPLICATION NUMBER: 60/117,531  
 ; PRIOR FILING DATE: 2000-02-01  
 ; PRIOR APPLICATION NUMBER: 60/117,844  
 ; PRIOR FILING DATE: 1999-01-29  
 ; PRIOR APPLICATION NUMBER: 60/118,206  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: 60/126,593  
 ; PRIOR FILING DATE: 1999-03-26  
 ; PRIOR APPLICATION NUMBER: 60/134,093  
 ; PRIOR FILING DATE: 1999-05-14  
 ; PRIOR APPLICATION NUMBER: 60/134,092  
 ; PRIOR FILING DATE: 1999-05-14  
 ; PRIOR APPLICATION NUMBER: 60/165,124  
 ; PRIOR FILING DATE: 1999-11-12  
 ; PRIOR APPLICATION NUMBER: 60/165,086  
 ; PRIOR FILING DATE: 1999-11-12  
 ; NUMBER OF SEQ ID NOS: 292  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 188  
 ; LENGTH: 335  
 ; TYPE: PRT  
 ; ORGANISM: Mycobacterium tuberculosis  
 ; US-09-712-363-188

Query Match 45.2%; Score 733; DB 10; Length 335;  
 Best Local Similarity 47.0%; Pred. No. 3,8e-63;  
 Matches 147; Conservative 66; Mismatches 92; Indels 8; Gaps 3;

Qy 2 QIVLANPFGAGVRAIEIVDAIEAFGAPLYRHEVNNRTVVDGLKQKAVFIEELS 61  
 Db 27 RVLAEPRGVCAGVRAVETVERALQKHPYVYHEIVNRRHVDTLAKAGAVFEETE 86  
 Qy 62 DVPVGSYLIFSANGSKEVQOAEEROLYFADATCPULVTKEQVAKAKGREFVLIGH 121  
 Db 87 QVPEGAIVVPSAIGVSPAVRAGADERGLQVADATCPULVAKVHEAARFPAARBDTWVF 146

Qy 122 AGHPEVGTMGQVEKCTEGGGIYLVETPEDEVRLKNNPNDLAVYVQTLSMTDTKVMD 181  
 Db 147 EGHPEVGTMGQVEKCTEGGGIYLVETPEDEVRLKNNPNDLAVYVQTLSMTDTKVMD 201  
 Qy 182 ALRBPSPSIXEOKKDDICVATONRQDAVHDLAKISDLIVGSPNSNSNRLRELVAVQJG 241  
 Db 202 RLRRFPKLDPPESDDICVATNRQALQSMWECVVLVIGSCNSNSRRLRELVAVQJG 261  
 Qy 242 -KQAVLIDYQDKODMLEGIEVVGVTAGASAPVLYOEVITDOLKAWGETTSVRENSGI 300  
 Db 262 ARAHLVMDADDIDGALMDGVTTVGTSGASVPEVLYRQVRLERLAECCGYDI--VQVPTTA 319  
 Qy 301 EEKVFSPKPELK 313  
 Db 320 NETLVFALPRELR 332

RESULT 8  
 US-10-289-762-1090  
 ; Sequence 1090, Application US/10289762  
 ; Publication No. US20040006218A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Griffls, R.  
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, frag  
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, i  
 ; FILE REFERENCE: 9710-003-999  
 ; CURRENT APPLICATION NUMBER: US/10/289,762  
 ; CURRENT FILING DATE: 2003-03-27  
 ; NUMBER OF SEQ ID NOS: 6849  
 ; SEQ ID NO 1090  
 ; LENGTH: 310  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia pneumoniae  
 ; US-10-289-762-1090

Query Match 40.9%; Score 663; DB 12; Length 310;  
 Best Local Similarity 45.0%; Pred. No. 2.4e-56;  
 Matches 143; Conservative 59; Mismatches 96; Indels 18; Gaps 3;

Qy 2 QIVLANPFGAGVRAIEIVDAIEAFGAPLYRHEVNNRTVVDGLKQKAVFIEELS 61  
 Db 3 KLILCNPRGSCGVRAIQVEVALLEKMGAPIYVKEHIVNRRHVNLRAKKAIFVEELV 62  
 Qy 62 DVPVGSYLIFSANGSKEVQOAEEROLYFADATCPULVTKEQVAKAKGREFVLIGH 121  
 Db 63 DVEGERVLYSAHGIPEVRAEAKRKLIDIDATGGLVTKVHNAKLYASKGYKILIGH 122  
 Qy 122 AGHPEVGTMGQVEKCTEGGGIYLVETPEDEVRLKNNPNDLAVYVQTLSMTDTKVMD 181  
 Db 123 KHVHVIIVIGEVPE-----HIVVEKVADEVALPFSDDPLFYITQTLSDIVCEISS 177  
 Qy 182 ALRBPSPSIXEOKKDDICVATONRQDAVHDLAKISDLIVGSPNSNSNRLRELVAVQJG 241  
 Db 178 ALIKRYPBIIITLPSISICVATYNNOKLRSVLSKVNYYVYVGVNNSNSNRLRELVAVQJG 237  
 Qy 242 KQAVLIDYQDKODMLEGIEVVGVTAGASAPVLYOEVITDOLKAWGETTSVRENSGI 301  
 Db 238 VADILNPNEDIDNIVVHSGDIATAGASTPEDEVQACIRKL-----SSLIPGQVE 290  
 Qy 302 -----EEKVFSPKPELK 313  
 Db 291 NDIPAVEDVVFQLPRELR 308

RESULT 9  
 US-10-022-832-36  
 ; Sequence 36, Application US/10022832  
 ; Publication No. US20030059896A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: COUTURE, FRANCE  
 ; APPLICANT: HAMEL, JOSEPH



Db 880 -----MESRNVGSGSKARDVLYK 900

RESULT 12  
US-09-712-363-282

Sequence 282, Application US/09712363  
Patent No. US20020164588A1

GENERAL INFORMATION:

APPLICANT: Eisendberg, David  
APPLICANT: Rotstein, Sergio H.

APPLICANT: Marcotte, Edward M.

TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS

FILE REFERENCE: 07419-032001

CURRENT APPLICATION NUMBER: US/09/712,363

CURRENT FILING DATE: 2000-11-13

PRIOR APPLICATION NUMBER: PCT/US00/02246

PRIOR FILING DATE: 2000-02-28

PRIOR APPLICATION NUMBER: 60/117,844

PRIOR FILING DATE: 1999-01-29

PRIOR APPLICATION NUMBER: 60/118,206

PRIOR FILING DATE: 1999-02-01

PRIOR APPLICATION NUMBER: 60/126,593

PRIOR FILING DATE: 1999-03-26

PRIOR APPLICATION NUMBER: 60/134,093

PRIOR FILING DATE: 1999-05-14

PRIOR APPLICATION NUMBER: 60/134,092

PRIOR FILING DATE: 1999-05-14

PRIOR APPLICATION NUMBER: 60/165,124

PRIOR FILING DATE: 1999-11-12

PRIOR APPLICATION NUMBER: 60/165,086

NUMBER OF SEQ ID NOS: 292

SOFTWARE: PASTESEQ for Windows Version 4.0

SEQ ID NO 282

LENGTH: 578

TYPE: PRT

ORGANISM: Mycobacterium tuberculosis

US-09-712-363-282

Query Match

Best Local Similarity 20.4%; Pred. No. 0.11;

Matches 78; Conservative 61; Mismatches 105; Indels 138; Gaps 18;

6 ANRPGFC-----AGVRAIRIVDQAIENFAPIVYRHEVYH 41

67 ANPGGVESCVSLAPNAPGSDIVVELDASHGVDDTRRLDR--AFYAPVQSRVYV-- 121

42 NRTVVDGLKQKGAVFIEBLSVPGSYLIFSAGVSK-----EVOQGEERQULVFPDA 94

122 -----FIVDEANHVTTAGFNALLKIVEEPHELIITFIPT 155

95 TCP-----LVTKVH-----MQUAKAKQGRVILLIGHAGPEVEGTMGQYE 135

156 TEPEKNTPTIRSRTHHPRFLPRTMRALLARICQ--EGVVVDDAVYLV----- 205

136 KCTEGGGIYIVTEPD--VENLKNNPNNDIAYVTO--TSMPTDTKMTVDALMEQFP 188

206 -IRAGGG-----SPRDLTSLVDQLLAGRAD--THVLYTRALGLGVDVALIDAV--D 254

189 SIKKQKODICVATQRQDAVHD-----LAKISLILVVGSPNSNS----- 230

255 ALAACDAALFGAIESVLDGCHDRPRPATDLERFRLIIVQSVEDASAGVVDAPEDAL 314

231 NRLREIVOLGKPAVLIIDYQDLKQDMLEGIEVVGTAASAPVIVQEVYIDOL--KAM 287

315 DRMEQARIGRAT--LTRYAVAVQAGL-----GEMKATAPRLILVVCARILLPSAS 366

286 GGETSVRENGSIEKVFISIP 309

367 DASALQORVERITRLDMSIP 388

RESULT 13

US-10-369-493-5415

Sequence 5415, Application US/10369493

Publication No. US20030233675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

PLANTS WITH IMPROVED PROPERTIES

FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 5415

LENGTH: 925

TYPE: PRT

ORGANISM: Caenorhabditis elegans

US-10-369-493-5415

Query Match

Best Local Similarity 23.1%; Pred. No. 0.23;

Matches 77; Conservative 53; Mismatches 134; Indels 70; Gaps 17;

3 IVLANPRGFCAGVDRATIEVDQAIENFAPIVYRHEVYHNRFTVVDGLKQKGAVFIEBLS 62

143 MILNNAKFTIEYRAQAEIHDVIT--VPVY--TPARLAVEBAQAGLTVQLIN 196

63 VPVGSYLIIFSAHGVSKVQGEAEERQUTVFD--ATCPVTVYRQVAKHAGQREVL 118

197 DGTAAAL--SHGIFRKSIGEKPRIMVYDMQAATTAITYFVKVKEVKEVQPKPTVL 253

119 -IGHAGHPEVGTMGQYKTEGGGTYLVETPEVNLKVN--NPNDLAYVYVOTILSMT 174

254 GVG-----PDRITGGIE-MTNRLRDLIMFE--KNYKTKMTNTRRA-MTKFKEAL 303

175 DTKVMDALNEQPSIEKQKDDICVATQRQDAVHDLAKISLILVVGSPNSNSNRLR 234

304 RLKQVLSANHEHFAQISABHDIDAKLVREDFNH--LISMESRFGSP-TEQALRMA 359

235 EIAVOLGKPAVLIIDYQDLKQDMLEGIEVVGTAASAPV--LVQEV----- 281

360 QIP-----DDIDQFVLMG-----AGTRVPRVQEVQKTIQTKETKELNT 400

282 DQLKANGGETTSVRENSG-----IEKVVFSI 308

401 DEAVAMGALLPQAHLSKGFVKPKNIEKVIIPV 434

RESULT 14

US-10-369-493-1052

Sequence 1052, Application US/10369493

Publication No. US20030233675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

PLANTS WITH IMPROVED PROPERTIES

FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 1052  
 LENGTH: 501  
 TYPE: PRT  
 ORGANISM: Methanococcus jannaschii  
 US-10-369-493-1052

Query Match 6.4%; Score 104.5; DB 12; Length 501;  
 Best Local Similarity 20.5%; Pred. No. 0.24;  
 Matches 69; Conservative 51; Mismatches 134; Indels 83; Gaps 15;

QY 14 GVDRAIRIVDQAI---EAFGAPITYR-----HEVWNR-RIVVDGLK 50  
 DB 133 GSEGAIEDIDBAIETABAIQFPVYVYKASGGGGMGMSVAVYSKEELKEVTEBARILAKSAF 192  
 QY 51 QKGAVFIEELSDVP--VGSYLIFSAHG-----VSKEVQQAEBERQUTVFDTCVLYKVN 103  
 DB 193 GDFVFIERYLLENPHIEIQLGDKHGNIIHLGDECSIQRRHQL-TERAPSPIMTE-- 249  
 QY 104 MQVAKHAKOGREVIILIGHGHEVEGTMGQYEKTEGGGIVLVEPDEVRMLAKVNPNDL 163  
 DB 250 ---EUREKGEAIIAGGAINYDSAGTV---EFLYENGNYFLE-----NMTRIQVHHT- 297  
 QY 164 AYTTLTSMYTDTKYVVDALREQPSIREKKDIDICATQ---NRQAVHDLAKISDLIL 220  
 DB 298 --VTGQVTGIDIVKAMIKIAGBELTLKQSDVKIRGHAIIECRINAEPLNDPVPQPKIX 355  
 QY 221 VVGSPPNSNSNRLEIAVQIGKPAVLIDTYODLKQDMLEGBVVGVTAGASAPVILQEV 280  
 DB 356 LTRSP-----GGPVRIDS-----GVYGAEIIP-YIDSM 384  
 QY 281 IDQLXAMG--ETTSVRENSGIEEKVFSIPKELKH 315  
 DB 385 IAKLITYGNSREEAIAFMKRALREYIIGVKTNIPEH 421

RESULT 15  
 US-10-369-493-19263  
 ; Sequence 19263, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianning  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052) B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; PRIOR FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 19263  
 ; LENGTH: 1130  
 ; TYPE: PRT  
 ; ORGANISM: Myxococcus xanthus  
 ; US-10-369-493-19263

Query Match 6.4%; Score 104.5; DB 12; Length 1130;  
 Best Local Similarity 16.9%; Pred. No. 0.86;  
 Matches 70; Conservative 64; Mismatches 124; Indels 155; Gaps 13;

QY 38 EYVHNRITVVDGLKQKGAVFIEELSDV-----EVSYLI-----FSA 73  
 DB 604 EYAEIADITINAMTQTLISPAQCVTVARIVGVGKLGAGAAVVPVAGATMKDFTNNVILA 663  
 QY 74 HGSVKEVQGEAE-----ERQUTV-----EDATCPVYTKVHM 104  
 DB 664 NMLTDQYRNIAEYTTAVAKGDLSKITVDAKGEVLELKSITNTMVDQLNSPAAEYTVAK 723  
 QY 105 QVAKHAKOGREVIILIGHGHEVEGTMGQYEKTEGGGIVLVEPDEVR----- 153

DB 724 EYGTGKLG-----GQAEVRGSGWMDLTDNNVFAVNLITQYRIVRYTAVAN 774  
 QY 154 -----NLKVNPNDLAIVTQTLISMTDT-----KVVVDALR-----EOPPSIKE 192  
 DB 775 GDLNQLTWDAGELIADITINAMTQTLISPAQCVTVARIVGVGKLGQAEVPGVAG 834  
 QY 193 QKRD---DICATQNRQDAVHDLAKIS-----DL-----ILVGSPPNSNSNR 232  
 DB 835 TKKDLTNNVNLANNLTAVYRNIAEYTTAVANGDLSKKITVDAKGEVLELKSITNTMVDQ 894  
 QY 233 LREIAVQIGKPAVLIDT-----YODLKQDMLEGBVVGVTAGASAPV----- 275  
 DB 895 LIAFAAEYTVAKGEVGTGKLGQADVHDLGVMMDLTDNNVLAGNLTDQVRNIAXYTT 954  
 QY 276 -----LVGEVINDLXAMGETTSVRENSGIEEKV 304  
 DB 955 AVANGDLSOKITVSYKGEVLELKSITNTMVDQLRAFASEYTVAKGEVGTGKLG 1007

Search completed: January 29, 2004, 16:21:19  
 Job time : 22.7073 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 15:54:19 ; Search time 10.7171 Seconds

(without alignments)  
3044.503 Million cell updates/sec

Title: US-09-941-947a-14

Perfect score: 806  
Sequence: 1 MIRVGMGYDVRHNDGHI.....LGFEGKKEGIAVQAVLIER 157

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

1: Published Applications AA:  
1: /cgn2\_6/prodata/1/pubpa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/prodata/1/pubpa/PC1\_NEW\_PUB.pep:\*  
3: /cgn2\_6/prodata/1/pubpa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/prodata/1/pubpa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/prodata/1/pubpa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/prodata/1/pubpa/PC1US\_PUBCOMB.pep:\*  
7: /cgn2\_6/prodata/1/pubpa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/prodata/1/pubpa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/prodata/1/pubpa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/prodata/1/pubpa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/prodata/1/pubpa/US09\_PUBCOMB.pep:\*  
12: /cgn2\_6/prodata/1/pubpa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/prodata/1/pubpa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/prodata/1/pubpa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/prodata/1/pubpa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/prodata/1/pubpa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/prodata/1/pubpa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/prodata/1/pubpa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	806	100.0	157	US-09-934-903-10
2	806	100.0	157	US-09-934-868-66
3	806	100.0	157	US-09-941-947a-14
4	496	61.5	170	US-10-174-410-12
5	360.5	44.7	223	US-10-859-194A-162
6	251	31.1	174	US-10-156-761-11502
7	247.5	30.7	158	US-10-128-713A-10
8	213.5	26.5	160	US-09-138-625-6421
9	153	19.0	32	US-10-174-410-135
10	153	19.0	199	US-10-289-762-585
11	143	17.7	32	US-10-174-410-129
12	138	17.1	32	US-10-174-410-133
13	138	17.1	32	US-10-174-410-134
14	138	17.1	32	US-10-174-410-136
15	138	17.1	32	US-10-174-410-149

15	133	16.5	32	15	US-10-174-410-125	Sequence 125, App
16	127	15.8	32	15	US-10-174-410-131	Sequence 131, App
17	126	15.6	32	15	US-10-174-410-127	Sequence 127, App
18	126	15.6	32	15	US-10-174-410-150	Sequence 150, App
19	125	15.5	32	15	US-10-174-410-128	Sequence 128, App
20	124	15.4	32	15	US-10-174-410-122	Sequence 122, App
21	124	15.4	32	15	US-10-174-410-140	Sequence 140, App
22	123	15.3	32	15	US-10-174-410-121	Sequence 121, App
23	123	15.3	32	15	US-10-174-410-124	Sequence 124, App
24	123	15.3	32	15	US-10-174-410-148	Sequence 148, App
25	119	14.8	32	15	US-10-174-410-132	Sequence 132, App
26	116	14.4	32	15	US-10-174-410-153	Sequence 153, App
27	116	14.4	32	15	US-10-174-410-130	Sequence 130, App
28	115	14.3	32	15	US-10-174-410-138	Sequence 138, App
29	115	14.3	32	15	US-10-174-410-154	Sequence 154, App
30	115	14.3	32	15	US-10-174-410-139	Sequence 139, App
31	114	14.1	32	15	US-10-174-410-142	Sequence 142, App
32	114	14.1	32	15	US-10-174-410-143	Sequence 143, App
33	114	14.1	32	15	US-10-174-410-156	Sequence 156, App
34	114	14.1	32	15	US-10-174-410-151	Sequence 151, App
35	112	13.9	32	15	US-10-174-410-123	Sequence 123, App
36	112	13.9	32	15	US-10-174-410-123	Sequence 123, App
37	110	13.6	32	15	US-10-174-410-123	Sequence 123, App
38	109	13.5	32	15	US-10-174-410-158	Sequence 158, App
39	109	13.5	32	15	US-10-174-410-229	Sequence 229, App
40	109	13.5	32	15	US-10-174-410-230	Sequence 230, App
41	108	13.4	32	15	US-10-174-410-189	Sequence 189, App
42	106	13.2	32	15	US-10-174-410-126	Sequence 126, App
43	106	13.2	32	15	US-10-174-410-146	Sequence 146, App
44	106	13.2	32	15	US-10-174-410-155	Sequence 155, App
45	106	13.2	32	15	US-10-174-410-173	Sequence 173, App

## ALIGNMENTS

RESULT 1  
US-09-934-903-10  
Sequence 10, Application US/0934903  
Patent No. US20020102690A1  
GENERAL INFORMATION:  
APPLICANT: Koffas, Mattheos  
APPLICANT: Odum, J. Martin  
APPLICANT: Schenzle, Andreas J.  
APPLICANT: No. US20020102690A1ton, Kelley C.  
APPLICANT: Tomb, Jean-Francois  
APPLICANT: Bouviere, Pierre  
APPLICANT: Picataggio, Stephen  
APPLICANT: Cheng, Qiong  
TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production  
FILE REFERENCE: C11646 US NA  
CURRENT APPLICATION NUMBER: US/09/934,903  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/229,907  
PRIOR FILING DATE: September 1, 2001  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 10  
LENGTH: 157  
TYPE: PRT  
ORGANISM: Methylobionas 16a  
FEATURE:  
OTHER INFORMATION: Amino acid sequences encoded by ORFs  
US-09-934-903-10  
Query Match 100.0%; Score 806; DB 10; Length 157;  
Best Local Similarity 100.0%; Pred. No. 3.5e-81;  
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MIRVGMGYDVRHNDGHIILGVKIPYKGLFAHSDGVVLAADALIGAAIDICK 60  
Db 1 MIRVGMGYDVRHNDGHIILGVKIPYKGLFAHSDGVVLAADALIGAAIDICK 60  
Qy 61 HPPDIDNFGADSRVLLRHVYGIYKEXYLVNADVTIIAQPKNLPHYKGRANIAD 120



```
Db 61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQPRLPHVPGMRANIAD 120
Qy 121 LETDVPFINVKATTTTEKLGFBGRKEGIAVQAVLIER 157
Db 121 LETDVPFINVKATTTTEKLGFBGRKEGIAVQAVLIER 157

RESULT 2
US-09-934-868-66
; Sequence 66, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENTRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: C11596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 66
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Methylomonas 16a
; OTHER INFORMATION: Amino acid sequences encoded by ISPF
US-09-934-868-66

Query Match 100.0%; Score 806; DB 10; Length 157;
Best Local Similarity 100.0%; Pred. No. 3.5e-81;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIRVGMGYDVHRFNDGDHIIILGGYKIPYKGLFASHSDDGVVLAALADALIGAAALGDIGK 60
Db 1 MIRVGMGYDVHRFNDGDHIIILGGYKIPYKGLFASHSDDGVVLAALADALIGAAALGDIGK 60
Qy 61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQPRLPHVPGMRANIAD 120
Db 61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQPRLPHVPGMRANIAD 120
Qy 61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQPRLPHVPGMRANIAD 120
Db 61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQPRLPHVPGMRANIAD 120
Qy 121 LETDVPFINVKATTTTEKLGFBGRKEGIAVQAVLIER 157
Db 121 LETDVPFINVKATTTTEKLGFBGRKEGIAVQAVLIER 157

RESULT 3
US-09-941-947A-14
; Sequence 14, Application US/09941947A
; Publication No. US2003003528A1
; GENERAL INFORMATION:
; APPLICANT: Brzostowicz, Patricia C.
; APPLICANT: Cheng, Qiong
; APPLICANT: Diccasimo, Deana J.
; APPLICANT: Koffas, Mattheos
; APPLICANT: Miller, Edward S. Jr.
; APPLICANT: Odom, J. Martin
; APPLICANT: Picataggio, Steve
; APPLICANT: Rouviere, Pierre B.
; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
; FILE REFERENCE: C11903 US NA
; CURRENT APPLICATION NUMBER: US/09/941,947A
; CURRENT FILING DATE: 2001-09-01
; PRIOR APPLICATION NUMBER: 60/229,907
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14

LENGTH: 157
; TYPE: PRT
; ORGANISM: Methylomonas 16a
US-09-941-947A-14

Query Match 100.0%; Score 806; DB 11; Length 157;
Best Local Similarity 100.0%; Pred. No. 3.5e-81;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIRVGMGYDVHRFNDGDHIIILGGYKIPYKGLFASHSDDGVVLAALADALIGAAALGDIGK 60
Db 1 MIRVGMGYDVHRFNDGDHIIILGGYKIPYKGLFASHSDDGVVLAALADALIGAAALGDIGK 60
Qy 61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQPRLPHVPGMRANIAD 120
Db 61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQPRLPHVPGMRANIAD 120
Qy 121 LETDVPFINVKATTTTEKLGFBGRKEGIAVQAVLIER 157
Db 121 LETDVPFINVKATTTTEKLGFBGRKEGIAVQAVLIER 157

RESULT 4
US-10-174-410-12
; Sequence 12, Application US/10174410
; Publication No. US2003007313A1
; GENERAL INFORMATION:
; APPLICANT: Louie, Gordon V.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Sander, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF
; TITLE OF INVENTION: 2C-METHYL-D-BRYTHRITOL, 2,4-CYCLODIPHOSPHATE SYNTHASE RECPs
; FILE REFERENCE: S2498200300
; CURRENT APPLICATION NUMBER: US/10/174,410
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/299,058
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 336
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 1, 69, 109, 115
; OTHER INFORMATION: Xaa = Selenomethionine
US-10-174-410-12

Query Match 61.5%; Score 496; DB 15; Length 170;
Best Local Similarity 61.8%; Pred. No. 7.7e-47;
Matches 97; Conservative 21; Mismatches 39; Indels 0; Gaps 0;

Qy 1 MIRVGMGYDVHRFNDGDHIIILGGYKIPYKGLFASHSDDGVVLAALADALIGAAALGDIGK 60
Db 3 LIRIGHGPDVHAFEDPPLIIGVEYHETGPIHSGDVALLHLDALIGAAALGDIGK 62
Qy 61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQPRLPHVPGMRANIAD 120
Db 63 LFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQPRLPHVPGMRANIAD 122
Qy 121 LETDVPFINVKATTTTEKLGFBGRKEGIAVQAVLIER 157
Db 123 LQCDIEOVNVLATTTTEKLGFTGDEGIAVQAVLIER 159

RESULT 5
US-10-259-194A-162
; Sequence 162, Application US/10259194A
; Publication No. US20040010815A1
; GENERAL INFORMATION:
; APPLICANT: Lange, Markus B.
```

```

; APPLICANT: Ghassemlan, Majid
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Rumiaki
; APPLICANT: Kuepfer, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Rieke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF PLANT GENES
; FILE REFERENCE: 70029-NP
; CURRENT APPLICATION NUMBER: US/10/259,194A
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,743
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 662
; SOFTWARE: Patentlist.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 162
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-259-194A-162

```

```

Query Match          44.7%; Score 360.5; DB 12; Length 223;
Best Local Similarity 46.8%; Pred. No. 1.1e-31;
Matches 73; Conservative 27; Mismatches 55; Indels 1; Gaps 1;

```

```

QY 3 RVGNGYDVRFPNDGDHIIIGGVKIPYKGLKSHSDGVVTHALADALGAAAGDICKF 62
DB 67 RVGNGYDVRFPNDGDHIIIGGVKIPYKGLKSHSDGVVTHALADALGAAAGDICKF 126
QY 63 PD-TDPNFKGADSVLRHRYGIVKEXGYKLVNADVTIIQAAPKMLPHVPGKRAMIAADL 121
DB 127 PELDPRKAGDSVFRKRAVKLHENGTEKLNADATLIIQKRKISPFKRTINSNLCDL 186
QY 122 ETDV--DFINVKATTTEKLGFGKREGIAVQAVVLI 157
DB 187 GADPSVYVLEKAKTEKYSIGENRSIAAHTVVLIR 222

```

## RESULT 6

```

US-10-156-761-11502
; Sequence 11502, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OKURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11502
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11502

```

```

Query Match          31.1%; Score 251; DB 15; Length 174;
Best Local Similarity 37.6%; Pred. No. 1e-19;
Matches 59; Conservative 23; Mismatches 67; Indels 8; Gaps 3;

```

```

QY 4 VGMGYDVRFPNDGDHIIIGGVKIPYKGLKSHSDGVVTHALADALGAAAGDICKF 62
DB 21 VGIQDTHAFEDGELMCAKLGKGEGLAGHSDADVVAHAAACNALFSAAGLDIQHF 80
QY 63 PTDNFKGADSVLRHRYGIVKEXGYKLVNADVTIIQAAPKMLPHVPGKRAMIAADL 122
DB 81 GIGREPNGAGAGVILLTAAATVYAAAGFTIGNVAVQVGRPKL-----GRRDEAQKL 135
QY 123 TDV--DFINVKATTTEKLGFGKREGIAVQAVVLI 157
DB 136 SDVAGAPVSVSAGATTDLGIFGFRGDEGLMAVATLVLR 172

```

## RESULT 7

```

US-10-128-713A-10
; Sequence 10, Application US/10128713A
; Publication No. US20030170847A1
; GENERAL INFORMATION:
; APPLICANT: Bramucci, Michael G
; TITLE OF INVENTION: Genes Involved in Isoprenoid Compound Production
; FILE REFERENCE: C-1788
; CURRENT APPLICATION NUMBER: US/10/128,713A
; PRIOR FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Rhodococcus erythropolis
US-10-128-713A-10

```

```

Query Match          30.7%; Score 247.5; DB 12; Length 158;
Best Local Similarity 37.8%; Pred. No. 2.2e-19;
Matches 59; Conservative 24; Mismatches 66; Indels 7; Gaps 2;

```

```

QY 2 IRVNGYDVRFPNDGDHIIIGGVKIPYKGLKSHSDGVVTHALADALGAAAGDICKF 61
DB 1 MRVGLGTDVHPVIEVGRCDMAAGLFEEDACSGSDSDVVAHLCALISAAGIDGVS 60
QY 62 PPDTPNFKGADSVLRHRYGIVKEXGYKLVNADVTIIQAAPKMLPHVPGKRAMIAADL 121
DB 61 FGTGRPEWDVSGARMLAEVRKLLHENGTFYGNAAVYIGNRKTL-----GPRDEAQKL 115
QY 122 ETDV--DFINVKATTTEKLGFGKREGIAVQAVVLI 155
DB 116 LSDILGAPVSVSATTDLGLGTRGEGIAAATLV 151

```

## RESULT 8

```

US-09-738-626-6421
; Sequence 6421, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16

```

; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO: 6421  
 ; LENGTH: 160  
 ; TYPE: PRT  
 ; ORGANISM: Corynebacterium glutamicum  
 ; US-09-738-626-6421

Query Match 26.5%; Score 213.5; DB 10; Length 160;  
 Best Local Similarity 32.9%; Pred. No. 1.3e-15;  
 Matches 51; Conservative 25; Mismatches 72; Indels 7; Gaps 2;

QY 3 RVGAGYDVHFRNDGHIILGGVKIPYKGLAEASDGDVVLHALLADALIGAAALGDIKHE 62  
 DB 8 RVGIATDAHQIEAGKPCWIALCLPEGYDGGGSHSDGVVAHAIYDALISASGLDGSFV 67

QY 63 PDIPTNFKGADSRVLRLHYVYIVKXGKLVNADVTIIAQPRLPHVPGKANIADLE 122  
 DB 68 GVGPEYDVGSGTOLKEVRELLSAHGVIIGNVAAQLVGGTPKF-----GPRREEAQVI 122

QY 123 TDV--DFINAKATTEKLGFEGRKEGIAVQAVALI 155  
 DB 123 SEIGAPCSLSATTTDHNKFTGRSEGRASVATNAV 157

RESULT 9  
 US-10-174-410-135  
 ; Sequence 135, Application US/10174410  
 ; Publication No. US20030072134A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Louie, Gordon V.  
 ; APPLICANT: Buchanan, Sean Grant  
 ; APPLICANT: Gajiwala, Ketan S.  
 ; APPLICANT: Sauder, J. Michael  
 ; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF  
 ; TITLE OF INVENTION: 2C-METHYL-D-ERYTHRITOL, 2,4-CYCLODIPHOSPHATE SYNTHASE MECPs  
 ; FILE REFERENCE: 524982000300  
 ; CURRENT APPLICATION NUMBER: US/10/174,410  
 ; CURRENT FILING DATE: 2002-06-17  
 ; PRIOR APPLICATION NUMBER: 60/299,058  
 ; PRIOR FILING DATE: 2001-06-18  
 ; NUMBER OF SEQ ID NOS: 336  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO: 135  
 ; LENGTH: 32  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-10-174-410-135

Query Match 19.0%; Score 153; DB 15; Length 32;  
 Best Local Similarity 87.5%; Pred. No. 7e-10;  
 Matches 28; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 40 VLHALLADALIGAAALGDIKHEPPTDPPNFKG 71  
 DB 1 VLHALLADALIGAAALGDIKHEPPTDPPNFKG 32

RESULT 10  
 US-10-289-762-585  
 ; Sequence 585, Application US/10289762  
 ; Publication No. US20040006218A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Louie, Gordon V.  
 ; APPLICANT: Buchanan, Sean Grant  
 ; APPLICANT: Gajiwala, Ketan S.  
 ; APPLICANT: Sauder, J. Michael  
 ; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF  
 ; TITLE OF INVENTION: 2C-METHYL-D-ERYTHRITOL, 2,4-CYCLODIPHOSPHATE SYNTHASE MECPs  
 ; FILE REFERENCE: 524982000300  
 ; CURRENT APPLICATION NUMBER: US/10/174,410  
 ; CURRENT FILING DATE: 2002-06-17  
 ; PRIOR APPLICATION NUMBER: 60/299,058

; CURRENT FILING DATE: 2003-03-27  
 ; NUMBER OF SEQ ID NOS: 6849  
 ; SEQ ID NO: 585  
 ; LENGTH: 199  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia pneumoniae  
 ; US-10-289-762-585

Query Match 19.0%; Score 153; DB 12; Length 199;  
 Best Local Similarity 29.4%; Pred. No. 8.5e-09;  
 Matches 47; Conservative 26; Mismatches 75; Indels 12; Gaps 5;

QY 3 RVGAGYDVHFRNDGHIILGGVKIPYKGLAEASDGDVVLHALLADALIGAAALGDIKHE 56  
 DB 40 RTGIGQDSHRLPSSSTYPCILGIIITDHCQFQANSDDIIFHAICNAISSYTNKILIG 99

QY 57 DIGKHFPPTDPPNFKG-ADSRVLRLHYVYIVKXGKLVNADVTIIAQPRLPHVPGMA 115  
 DB 100 KVADELDTQ---RGITDSGIYLEALKSLK-PNOKISHVALITEGSRPFLCTLSAIRQ 154

QY 116 NIADLEFDVDFINAKATTEKLGFEGRKEGIAVQAVALI 155  
 DB 155 NIAGVNLTPTDIGITATISGEGLSDFGQSGVCCFVLTV 194

RESULT 11  
 US-10-174-410-129  
 ; Sequence 129, Application US/10174410  
 ; Publication No. US20030072134A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Louie, Gordon V.  
 ; APPLICANT: Buchanan, Sean Grant  
 ; APPLICANT: Gajiwala, Ketan S.  
 ; APPLICANT: Sauder, J. Michael  
 ; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF  
 ; TITLE OF INVENTION: 2C-METHYL-D-ERYTHRITOL, 2,4-CYCLODIPHOSPHATE SYNTHASE MECPs  
 ; FILE REFERENCE: 524982000300  
 ; CURRENT APPLICATION NUMBER: US/10/174,410  
 ; CURRENT FILING DATE: 2002-06-17  
 ; PRIOR APPLICATION NUMBER: 60/299,058  
 ; PRIOR FILING DATE: 2001-06-18  
 ; NUMBER OF SEQ ID NOS: 336  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO: 129  
 ; LENGTH: 32  
 ; TYPE: PRT  
 ; ORGANISM: C. perfringens  
 ; US-10-174-410-129

Query Match 17.7%; Score 143; DB 15; Length 32;  
 Best Local Similarity 78.1%; Pred. No. 8.9e-09;  
 Matches 25; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 40 VLHALLADALIGAAALGDIKHEPPTDPPNFKG 71  
 DB 1 VLHALLADALIGAAALGDIKHEPPTDPPNFKG 32

RESULT 12  
 US-10-174-410-133  
 ; Sequence 133, Application US/10174410  
 ; Publication No. US20030072134A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Louie, Gordon V.  
 ; APPLICANT: Buchanan, Sean Grant  
 ; APPLICANT: Gajiwala, Ketan S.  
 ; APPLICANT: Sauder, J. Michael  
 ; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF  
 ; TITLE OF INVENTION: 2C-METHYL-D-ERYTHRITOL, 2,4-CYCLODIPHOSPHATE SYNTHASE MECPs  
 ; FILE REFERENCE: 524982000300  
 ; CURRENT APPLICATION NUMBER: US/10/174,410  
 ; CURRENT FILING DATE: 2002-06-17  
 ; PRIOR APPLICATION NUMBER: 60/299,058

PRIOR FILING DATE: 2001-06-18  
NUMBER OF SEQ ID NOS: 336  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 133  
LENGTH: 32  
TYPE: PRT  
ORGANISM: S. typhimurium  
US-10-174-410-133

Query Match 17.1%; Score 138; DB 15; Length 32;  
Best Local Similarity 84.4%; Pred. No. 3.2e-08;  
Matches 27; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 40 VTHALADAILGAAALGDIKGFPTDPNFKG 71  
DB 1 VALHALTDALGAAALGDIKGFPTDPAFKG 32

## RESULT 13

US-10-174-410-134  
Sequence 134, Application US/10174410  
Publication No. US20030073134A1  
GENERAL INFORMATION:  
APPLICANT: Louie, Gordon V.  
APPLICANT: Buchanan, Sean Grant  
APPLICANT: Gajiwala, Ketan S.  
APPLICANT: Sauder, J. Michael  
TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF  
FILE REFERENCE: 524982000300  
CURRENT FILING DATE: 2002-06-17  
PRIOR FILING DATE: 2001-06-18  
NUMBER OF SEQ ID NOS: 336  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 134  
LENGTH: 32  
TYPE: PRT  
ORGANISM: S. enterica  
US-10-174-410-134

Query Match 17.1%; Score 138; DB 15; Length 32;  
Best Local Similarity 84.4%; Pred. No. 3.2e-08;  
Matches 27; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 40 VTHALADAILGAAALGDIKGFPTDPNFKG 71  
DB 1 VALHALTDALGAAALGDIKGFPTDPAFKG 32

## RESULT 14

US-10-174-410-136  
Sequence 136, Application US/10174410  
Publication No. US20030073134A1  
GENERAL INFORMATION:  
APPLICANT: Louie, Gordon V.  
APPLICANT: Buchanan, Sean Grant  
APPLICANT: Gajiwala, Ketan S.  
APPLICANT: Sauder, J. Michael  
TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF  
FILE REFERENCE: 524982000300  
CURRENT FILING DATE: 2002-06-17  
PRIOR FILING DATE: 2001-06-18  
NUMBER OF SEQ ID NOS: 336  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 136  
LENGTH: 32  
TYPE: PRT  
ORGANISM: Escherichia coli

US-10-174-410-136

Query Match 17.1%; Score 138; DB 15; Length 32;  
Best Local Similarity 84.4%; Pred. No. 3.2e-08;  
Matches 27; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 40 VTHALADAILGAAALGDIKGFPTDPNFKG 71  
DB 1 VALHALTDALGAAALGDIKGFPTDPAFKG 32

## RESULT 15

US-10-174-410-149  
Sequence 149, Application US/10174410  
Publication No. US20030073134A1  
GENERAL INFORMATION:  
APPLICANT: Louie, Gordon V.  
APPLICANT: Buchanan, Sean Grant  
APPLICANT: Gajiwala, Ketan S.  
APPLICANT: Sauder, J. Michael  
TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF  
FILE REFERENCE: 524982000300  
CURRENT FILING DATE: 2002-06-17  
PRIOR FILING DATE: 2001-06-18  
NUMBER OF SEQ ID NOS: 336  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 149  
LENGTH: 32  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-10-174-410-149

Query Match 17.1%; Score 138; DB 15; Length 32;  
Best Local Similarity 84.4%; Pred. No. 3.2e-08;  
Matches 27; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 40 VTHALADAILGAAALGDIKGFPTDPNFKG 71  
DB 1 VALHALTDALGAAALGDIKGFPTDPAFKG 32

Search completed: January 29, 2004, 16:21:17  
Job time : 10.7171 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM protein - protein search, using BW model

Run on: January 29, 2004, 15:54:19 ; Search time 37.1146 Seconds  
(without alignments)  
3044.503 Million cell updates/sec

Title: US-09-941-947a-16

Perfect score: 2788  
Sequence: 1 MTRKFFITGVSSLGKIA.....HALSGFVEAAAKKTGCTA 544

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues  
Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2788	100.0	544	10	US-09-934-903-12
2	2788	100.0	544	10	US-09-934-868-70
3	2788	100.0	544	11	US-09-941-947a-16
4	2008	72.0	544	9	US-09-815-242-11933
5	1816	68.7	545	9	US-09-815-242-10725
6	1908.5	68.0	545	9	US-09-815-242-11175
7	1783	64.0	511	9	US-09-815-242-13865
8	1576	56.5	536	9	US-09-815-242-12372
9	1576	56.5	536	9	US-09-815-242-12978
10	1569	56.3	536	9	US-09-815-242-4984
11	1569	56.3	536	9	US-09-815-242-10727
12	1561	56.0	535	9	US-09-815-242-13376
13	1444.5	51.8	538	9	US-09-815-242-11320
14	1432.5	51.4	549	15	US-10-156-761-14031
15	1372	49.2	554	10	US-09-738-626-5066

16	1355.5	48.6	537	12	US-10-289-762-246	Sequence 246, App
17	1006.5	36.1	1003	12	US-10-161-051-91	Sequence 91, Appl
18	513	18.4	298	12	US-10-264-049-2881	Sequence 2881, Ap
19	353.5	12.7	199	15	US-10-106-698-4980	Sequence 4980, Ap
20	345.5	12.4	95	9	US-09-925-297-773	Sequence 773, App
21	327.5	11.7	194	12	US-09-864-408A-3914	Sequence 3914, Ap
22	245	8.8	59	12	US-09-864-408A-2350	Sequence 2350, Ap
23	204	7.3	135	12	US-10-264-049-4217	Sequence 4217, Ap
24	126.5	4.5	238	12	US-10-314-657-14	Sequence 34, Appl
25	126	4.5	374	12	US-10-369-493-10131	Sequence 10131, A
26	125.5	4.5	1390	12	US-10-369-493-11073	Sequence 11073, A
27	124	4.4	374	12	US-10-369-493-15844	Sequence 15844, A
28	124	4.4	374	12	US-10-369-493-16223	Sequence 16223, A
29	123.5	4.4	371	12	US-10-369-493-1338	Sequence 29, Appl
30	122	4.4	845	10	US-09-738-626-6347	Sequence 4347, Ap
31	120	4.3	382	12	US-10-369-493-691	Sequence 691, App
32	120	4.3	382	12	US-10-369-493-15474	Sequence 15474, A
33	120	4.3	382	16	US-10-210-115-33	Sequence 33, Appl
34	116.5	4.2	656	12	US-10-369-493-18670	Sequence 18670, A
35	115.5	4.1	650	12	US-10-369-493-1338	Sequence 1338, Ap
36	115.5	4.1	650	12	US-10-369-493-20401	Sequence 20401, A
37	115.5	4.1	1687	14	US-10-094-679-3	Sequence 3, Appl1
38	114	4.1	1071	12	US-10-369-493-23117	Sequence 23117, A
39	113.5	4.1	106	12	US-09-864-408A-2256	Sequence 2256, Ap
40	113	4.1	362	12	US-10-369-493-17332	Sequence 17332, A
41	112.5	4.0	371	12	US-10-369-493-11567	Sequence 11567, A
42	112.5	4.0	537	12	US-10-369-493-89	Sequence 89, Appl
43	111.5	4.0	728	12	US-10-369-493-15429	Sequence 15429, A
44	111.5	4.0	728	12	US-10-369-493-15797	Sequence 15797, A
45	111.5	4.0	728	12	US-10-369-493-16179	Sequence 16179, A

## ALIGNMENTS

RESULT 1  
US-09-934-903-12  
Sequence 12, Application US/09934903  
Patient No. US20020102690A1  
GENERAL INFORMATION:  
APPLICANT: Kofias, Matheos  
APPLICANT: Odum, J. Martin  
APPLICANT: Schenzle, Andreas J.  
APPLICANT: No. US20020102690A1on, Kelley C.  
APPLICANT: Tomb, Jean-Francois  
APPLICANT: Rouviere, Pierre  
APPLICANT: Picataggio, Stephen  
TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production  
FILE REFERENCE: C14646 US NA  
CURRENT APPLICATION NUMBER: US/09/934,903  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/229,907  
PRIOR FILING DATE: September 1, 2001  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 12  
LENGTH: 544  
TYPE: PRP  
ORGANISM: Methylobionas 16a  
FEATURE:  
OTHER INFORMATION: Amino acid sequences encoded by ORF6  
US-09-934-903-12  
Query Match 100.0%; Score 2788; DB 10; Length 544;  
Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MTRKFFITGVSSLGKIAASLAIEDRGKVTITTKDDPYINVDPGTMSPPQGEVF 60  
DB 1 MTRKFFITGVSSLGKIAASLAIEDRGKVTITTKDDPYINVDPGTMSPPQGEVF 60  
QY 61 VTDEGAETDLDLGHYRFLKTTMKKNNFTTQGYEVLNERRGDTGATGYIPIITD 120

Db 61 VTEDGAEFDLGLGHERFLKTTMTKNNFTTGQYVEQLRNERGDLGATVQYIPHTD 120  
QY 121 EIKRRVYSABGKVALLEVGGTVGDIESTLPFTETIRQMGVELGRDALFIHLTVYIK 180  
Db 121 EIKRRVYSABGKVALLEVGGTVGDIESTLPFTETIRQMGVELGRDALFIHLTVYIK 180  
QY 181 SAGELKTPQTHSVYELTTIGIOPILICRSEOPIPASERKIALFTNVAEKAVISAIDA 240  
Db 181 SAGELKTPQTHSVYELTTIGIOPILICRSEOPIPASERKIALFTNVAEKAVISAIDA 240  
QY 241 DTIYRIPLLREOGDLVDVQRLDVPADLSAMEKVVDGLTHPTDEVSIAIVGYVDH 300  
Db 241 DTIYRIPLLREOGDLVDVQRLDVPADLSAMEKVVDGLTHPTDEVSIAIVGYVDH 300  
QY 301 TDVAYSALBEALIHAGIHRHAKVQISYIDSETIEBGRPAKKNVDAILVPGFGRGVEGK 360  
Db 301 TDVAYSALBEALIHAGIHRHAKVQISYIDSETIEBGRPAKKNVDAILVPGFGRGVEGK 360  
QY 361 ISTVFAFENKIPYIGICLGMQSAVIEPARNVGLBGHSTFPLPKSPHPVIGLITEMD 420  
Db 361 ISTVFAFENKIPYIGICLGMQSAVIEPARNVGLBGHSTFPLPKSPHPVIGLITEMD 420  
QY 421 EAGELVTDESDSLGCTRLGAOKCRLKADSLAFOLYOKVITERHHRHREYFNNQYIKOL 480  
Db 421 EAGELVTDESDSLGCTRLGAOKCRLKADSLAFOLYOKVITERHHRHREYFNNQYIKOL 480  
QY 481 EAAGKPFSGKSLDGLVLEIIELEPHWFLACQFHPFTSTPRNGHALPSGFVEAAAKHT 540  
Db 481 EAAGKPFSGKSLDGLVLEIIELEPHWFLACQFHPFTSTPRNGHALPSGFVEAAAKHT 540  
QY 541 OGTA 544  
Db 541 OGTA 544

RESULT 2  
US-09-934-868-70  
Sequence 70, Application US/09934868  
Patent No. US20020137190A1  
GENERAL INFORMATION:  
APPLICANT: Koffas, Mattheos  
APPLICANT: Odum, James M.  
APPLICANT: Schenzle, Andreas J.  
TITLE OF INVENTION: IDENTIFYING METHANOTROPHIC BACTERIAL STRAIN  
FILE REFERENCE: CIL1596 US NA  
CURRENT APPLICATION NUMBER: US/09/934, 868  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/229, 858  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 70  
LENGTH: 544  
TYPE: PRF  
ORGANISM: Methylomonas 16a  
FEATURE:  
OTHER INFORMATION: Amino acid sequences encoded by ORF6 - PYRG  
US-09-934-868-70

Query Match 100.0%; Score 2788; DB 10; Length 544;  
Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKPFIITGVVSSIGKGIASLSAAILDEBGLKVTITKDDPYINVPGMTSPQHGVEF 60  
Db 1 MTKPFIITGVVSSIGKGIASLSAAILDEBGLKVTITKDDPYINVPGMTSPQHGVEF 60  
QY 61 VTEDGAEFDLGLGHERFLKTTMTKNNFTTGQYVEQLRNERGDLGATVQYIPHTD 120  
Db 61 VTEDGAEFDLGLGHERFLKTTMTKNNFTTGQYVEQLRNERGDLGATVQYIPHTD 120  
QY 121 EIKRRVYSABGKVALLEVGGTVGDIESTLPFTETIRQMGVELGRDALFIHLTVYIK 180

Db 121 EIKRRVYSABGKVALLEVGGTVGDIESTLPFTETIRQMGVELGRDALFIHLTVYIK 180  
QY 181 SAGELKTPQTHSVYELTTIGIOPILICRSEOPIPASERKIALFTNVAEKAVISAIDA 240  
Db 181 SAGELKTPQTHSVYELTTIGIOPILICRSEOPIPASERKIALFTNVAEKAVISAIDA 240  
QY 241 DTIYRIPLLREOGDLVDVQRLDVPADLSAMEKVVDGLTHPTDEVSIAIVGYVDH 300  
Db 241 DTIYRIPLLREOGDLVDVQRLDVPADLSAMEKVVDGLTHPTDEVSIAIVGYVDH 300  
QY 301 TDVAYSALBEALIHAGIHRHAKVQISYIDSETIEBGRPAKKNVDAILVPGFGRGVEGK 360  
Db 301 TDVAYSALBEALIHAGIHRHAKVQISYIDSETIEBGRPAKKNVDAILVPGFGRGVEGK 360  
QY 361 ISTVFAFENKIPYIGICLGMQSAVIEPARNVGLBGHSTFPLPKSPHPVIGLITEMD 420  
Db 361 ISTVFAFENKIPYIGICLGMQSAVIEPARNVGLBGHSTFPLPKSPHPVIGLITEMD 420  
QY 421 EAGELVTDESDSLGCTRLGAOKCRLKADSLAFOLYOKVITERHHRHREYFNNQYIKOL 480  
Db 421 EAGELVTDESDSLGCTRLGAOKCRLKADSLAFOLYOKVITERHHRHREYFNNQYIKOL 480  
QY 481 EAAGKPFSGKSLDGLVLEIIELEPHWFLACQFHPFTSTPRNGHALPSGFVEAAAKHT 540  
Db 481 EAAGKPFSGKSLDGLVLEIIELEPHWFLACQFHPFTSTPRNGHALPSGFVEAAAKHT 540  
QY 541 OGTA 544  
Db 541 OGTA 544

RESULT 3  
US-09-941-947a-16  
Sequence 16, Application US/09941947a  
Publication No. US20030003528A1  
GENERAL INFORMATION:  
APPLICANT: Brzostowicz, Patricia C.  
APPLICANT: Cheng, Qiong  
APPLICANT: DiCosimo, Deana J.  
APPLICANT: Koffas, Mattheos  
APPLICANT: Miller, Edward S. Jr.  
APPLICANT: Odum, J. Martin  
APPLICANT: Picataggio, Steve  
TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE  
FILE REFERENCE: CIL1903 US NA  
CURRENT APPLICATION NUMBER: US/09/941, 947a  
CURRENT FILING DATE: 2001-09-01  
PRIOR APPLICATION NUMBER: 60/229, 907  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 60  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 16  
LENGTH: 544  
TYPE: PRF  
ORGANISM: Methylomonas 16a  
US-09-941-947a-16

Query Match 100.0%; Score 2788; DB 11; Length 544;  
Best Local Similarity 100.0%; Pred. No. 1.5e-254;  
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKPFIITGVVSSIGKGIASLSAAILDEBGLKVTITKDDPYINVPGMTSPQHGVEF 60  
Db 1 MTKPFIITGVVSSIGKGIASLSAAILDEBGLKVTITKDDPYINVPGMTSPQHGVEF 60  
QY 61 VTEDGAEFDLGLGHERFLKTTMTKNNFTTGQYVEQLRNERGDLGATVQYIPHTD 120  
Db 61 VTEDGAEFDLGLGHERFLKTTMTKNNFTTGQYVEQLRNERGDLGATVQYIPHTD 120

```

QY 121 EIKRRYESABEKDVALIEVGSTGNDISLPLFTIRMGVHLCGRDLFIHLTLVPIK 180
DB 121 EIKRRYESABEKDVALIEVGSTGNDISLPLFTIRMGVHLCGRDLFIHLTLVPIK 180
QY 181 SAGELTKPTQHSVKELEATIGIQPDLICRSEQPIPASERRKIALFTVNAEKAVISAIDA 240
DB 181 SAGELTKPTQHSVKELEATIGIQPDLICRSEQPIPASERRKIALFTVNAEKAVISAIDA 240
QY 241 DITIRIPVLAHQGLDVLVQDLRLDVPADLSANEKVYDGLTHEPTDEVSIAIVGYVDH 300
DB 241 DITIRIPVLAHQGLDVLVQDLRLDVPADLSANEKVYDGLTHEPTDEVSIAIVGYVDH 300
QY 301 TAYVSLNEALIHAGIHTRHVKQIISYIDSETTEAGTAKLKNVVALIVPGGFGERGVEGK 360
DB 301 TAYVSLNEALIHAGIHTRHVKQIISYIDSETTEAGTAKLKNVVALIVPGGFGERGVEGK 360
QY 361 ISTVPRARENKIPIYIGICLGMQSAVIEPARNVVGLGASHTEFLPKSPHPIVGLITEMD 420
DB 361 ISTVPRARENKIPIYIGICLGMQSAVIEPARNVVGLGASHTEFLPKSPHPIVGLITEMD 420
QY 421 EAGELVTREDESDLGTRRLGOKRCLKADSLAPQLYOKDVTERRHRYEENNOYKOL 480
DB 421 EAGELVTREDESDLGTRRLGOKRCLKADSLAPQLYOKDVTERRHRYEENNOYKOL 480
QY 481 EAAKMKFSGKSLDGLVLEIIELEPHPMFLACQFHEFTSTPRNGHALFSGFVBAARHKT 540
DB 481 EAAKMKFSGKSLDGLVLEIIELEPHPMFLACQFHEFTSTPRNGHALFSGFVBAARHKT 540
QY 541 OGTA 544
DB 541 OGTA 544

```

## RESULT 4

```

US-09-815-242-11933
Sequence 11933, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815.242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11933
LENGTH: 542
TYPE: PRF
ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11933

```

Query Match 72.0%; Score 2008; DB 9; Length 542;  
 Best Local Similarity 71.2%; Pred. No. 8,5e-181;  
 Matches 383; Conservative 64; Mismatches 91; Indels 0; Gaps 0;

```

QY 1 MRFPIITGVSSTIGKGIASGLAAIEDRGAKYITRKLDPIYINWDPGTMSPFQGVYF 60
DB 1 MRFPIITGVSSTIGKGIASGLAAIEDRGAKYITRKLDPIYINWDPGTMSPFQGVYF 60
QY 61 VTDEGAEITDLGHTERFLKTTMTKKNFTTQGVYEOYLRNKKDGYLGATVQVIPHITD 120
DB 61 VTDEGAEITDLGHTERFLKTTMTKKNFTTQGVYEOYLRNKKDGYLGATVQVIPHITD 120
QY 121 EIKRRYESABEKDVALIEVGSTGNDISLPLFTIRMGVHLCGRDLFIHLTLVPIK 180
DB 121 EIKRRYESABEKDVALIEVGSTGNDISLPLFTIRMGVHLCGRDLFIHLTLVPIK 180
QY 181 SAGELTKPTQHSVKELEATIGIQPDLICRSEQPIPASERRKIALFTVNAEKAVISAIDA 240
DB 181 SAGELTKPTQHSVKELEATIGIQPDLICRSEQPIPASERRKIALFTVNAEKAVISAIDA 240
QY 241 DITIRIPVLAHQGLDVLVQDLRLDVPADLSANEKVYDGLTHEPTDEVSIAIVGYVDH 300
DB 241 DITIRIPVLAHQGLDVLVQDLRLDVPADLSANEKVYDGLTHEPTDEVSIAIVGYVDH 300
QY 301 TAYVSLNEALIHAGIHTRHVKQIISYIDSETTEAGTAKLKNVVALIVPGGFGERGVEGK 360
DB 301 TAYVSLNEALIHAGIHTRHVKQIISYIDSETTEAGTAKLKNVVALIVPGGFGERGVEGK 360
QY 361 ISTVPRARENKIPIYIGICLGMQSAVIEPARNVVGLGASHTEFLPKSPHPIVGLITEMD 420
DB 361 ISTVPRARENKIPIYIGICLGMQSAVIEPARNVVGLGASHTEFLPKSPHPIVGLITEMD 420
QY 421 EAGELVTREDESDLGTRRLGOKRCLKADSLAPQLYOKDVTERRHRYEENNOYKOL 480
DB 421 EAGELVTREDESDLGTRRLGOKRCLKADSLAPQLYOKDVTERRHRYEENNOYKOL 480
QY 481 EAAKMKFSGKSLDGLVLEIIELEPHPMFLACQFHEFTSTPRNGHALFSGFVBAARHKT 538
DB 481 EAAKMKFSGKSLDGLVLEIIELEPHPMFLACQFHEFTSTPRNGHALFSGFVBAARHKT 538
QY 538 EAGELTKPTQHSVKELEATIGIQPDLICRSEQPIPASERRKIALFTVNAEKAVISAIDA 600
DB 538 EAGELTKPTQHSVKELEATIGIQPDLICRSEQPIPASERRKIALFTVNAEKAVISAIDA 600

```

## RESULT 5

```

US-09-815-242-10275
Sequence 10275, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815.242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16

```

NUMBER OF SEQ ID NOS: 14110  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO: 10275  
 LENGTH: 545  
 TYPE: PRT  
 ORGANISM: Escherichia coli  
 US-09-815-242-10275

Query Match 68.7%; Score 1916; DB 9; Length 545;  
 Best Local Similarity 67.7%; Pred. No. 4.3e-172;  
 Matches 363; Conservative 71; Mismatches 102; Indels 0; Gaps 0;

QY 2 TKTFITGGVSSISGKGLAASLALIEDGKATITKLDPIYNVDPGTMSPROHGEVY 61  
 DB 3 TNYIFVGGVSSISGKGLAASLALIEDGKATITKLDPIYNVDPGTMSPROHGEVY 62  
 QY 62 TEDGAEFDLDGHERERFLKTTMTKNNFTTGOVVEQVLRNERKQDYAGATVOYIPIHTDE 121  
 DB 63 TEDGAEFDLDGHERERFLKTTMTKNNFTTGOVVEQVLRNERKQDYAGATVOYIPIHTDE 122  
 QY 122 IKRKYBSAEKQVALLIEVGCTVDDISLPLETIRQXVBLGRDALPHTLTVPIYKS 181  
 DB 123 IKRKYBSAEKQVALLIEVGCTVDDISLPLETIRQXVBLGRDALPHTLTVPIYMA 182  
 QY 182 AGELTKETQSVKELRTIGIQPILICRSOPIPASERRKIALFTVAEKAVISAIDAD 241  
 DB 183 SGVYKTKTQSVKELRTIGIQPILICRSOPIPASERRKIALFTVAEKAVISAIDAD 242  
 QY 242 TTYRIPLLRQGLDVLVDQLRDLVPAADISANEKYVDGTHPTDVSIAIVKTYDHT 301  
 DB 243 SIYKIPALKSQGDDVFCERFLTCPEADLTMEQVLYKQANPVGCVTTIGWVKTELP 302  
 QY 302 DAVSLNKLTHAGIHPRKXQIYIDSETEAGTKLKNVDALIVPGGFGSGVEGKI 361  
 DB 303 DAVSLNKLTHAGIHPRKXQIYIDSETEAGTKLKNVDALIVPGGFGSGVEGKI 362  
 QY 362 STVPARENKIPYLGICLQWQSAVIEPARNVGLGASHTEFLPKSHPYVIGLITTEMD 421  
 DB 363 STVPARENKIPYLGICLQWQSAVIEPARNVGLGASHTEFLPKSHPYVIGLITTEMD 422  
 QY 422 AGEVYTRDESDIGGTWELGAKOCLKADSLAPOLYKQDVTEHRRHRYEFNNQYKOLE 481  
 DB 423 AGEVYTRDESDIGGTWELGAKOCLKADSLAPOLYKQDVTEHRRHRYEFNNQYKOLE 482  
 QY 482 AAGMFGSKSLDGLVLEIIELEPHWPLACQHPHEFTSTPRNGHALFSGEVEAAK 537  
 DB 483 AAGMFGSKSLDGLVLEIIELEPHWPLACQHPHEFTSTPRNGHALFSGEVEAAK 538

RESULT 6  
 US-09-815-242-11175  
 Sequence 11175, Application US/09815242  
 Patent No. US2002061569A1  
 GENERAL INFORMATION:  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari L.  
 APPLICANT: Zyskind, Judith W.  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John D.  
 APPLICANT: Carr, Grant J.  
 APPLICANT: Yamamoto, Robert T.  
 APPLICANT: Xu, H. Howard  
 TITLE OF INVENTION: Identification of Essential Genes in  
 TITLE OF INVENTION: Prokaryotes  
 FILE REFERENCE: ELITRA.011A  
 CURRENT APPLICATION NUMBER: US/09/815.242  
 PRIOR FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 NUMBER OF SEQ ID NOS: 14110  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO: 11175  
 LENGTH: 545  
 TYPE: PRT  
 ORGANISM: Haemophilus influenzae  
 US-09-815-242-11175

Query Match 68.5%; Score 1908.5; DB 9; Length 545;  
 Best Local Similarity 67.5%; Pred. No. 2.2e-171;  
 Matches 364; Conservative 71; Mismatches 103; Indels 1; Gaps 1;

QY 2 TKTFITGGVSSISGKGLAASLALIEDGKATITKLDPIYNVDPGTMSPROHGEVY 61  
 DB 3 TNYIFVGGVSSISGKGLAASLALIEDGKATITKLDPIYNVDPGTMSPROHGEVY 62  
 QY 62 TEDGAEFDLDGHERERFLKTTMTKNNFTTGOVVEQVLRNERKQDYAGATVOYIPIHTDE 121  
 DB 63 TEDGAEFDLDGHERERFLKTTMTKNNFTTGOVVEQVLRNERKQDYAGATVOYIPIHTDE 122  
 QY 122 IKRKYBSAEKQVALLIEVGCTVDDISLPLETIRQXVBLGRDALPHTLTVPIYKS 181  
 DB 123 IKRKYBSAEKQVALLIEVGCTVDDISLPLETIRQXVBLGRDALPHTLTVPIYPT 182  
 QY 182 AGELTKETQSVKELRTIGIQPILICRSOPIPASERRKIALFTVAEKAVISAIDAD 241  
 DB 183 AGELTKETQSVKELRTIGIQPILICRSOPIPASERRKIALFTVAEKAVISAIDAD 242  
 QY 242 TTYRIPLLRQGLDVLVDQLRDLVPAADISANEKYVDGTHPTDVSIAIVKTYDHT 301  
 DB 243 SIYKIPALKSQGDDVFCERFLTCPEADLTMEQVLYKQANPVGCVTTIGWVKTELP 302  
 QY 302 DAVSLNKLTHAGIHPRKXQIYIDSETEAGTKLKNVDALIVPGGFGSGVEGKI 361  
 DB 303 DAVSLNKLTHAGIHPRKXQIYIDSETEAGTKLKNVDALIVPGGFGSGVEGKI 362  
 QY 362 STVPARENKIPYLGICLQWQSAVIEPARNVGLGASHTEFLPKSHPYVIGLITTEMD 421  
 DB 363 STVPARENKIPYLGICLQWQSAVIEPARNVGLGASHTEFLPKSHPYVIGLITTEMD 422  
 QY 422 AGEVYTRDESDIGGTWELGAKOCLKADSLAPOLYKQDVTEHRRHRYEFNNQYKOLE 481  
 DB 423 AGEVYTRDESDIGGTWELGAKOCLKADSLAPOLYKQDVTEHRRHRYEFNNQYKOLE 482  
 QY 482 AAGMFGSKSLDGLVLEIIELEPHWPLACQHPHEFTSTPRNGHALFSGEVEAA-AHK 539  
 DB 483 AAGMFGSKSLDGLVLEIIELEPHWPLACQHPHEFTSTPRNGHALFSGEVEAA-AHK 541

RESULT 7  
 US-09-815-242-11865  
 Sequence 11865, Application US/09815242  
 Patent No. US2002061569A1  
 GENERAL INFORMATION:  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari L.  
 APPLICANT: Zyskind, Judith W.  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John D.  
 APPLICANT: Carr, Grant J.  
 APPLICANT: Yamamoto, Robert T.  
 APPLICANT: Xu, H. Howard  
 TITLE OF INVENTION: Identification of Essential Genes in  
 TITLE OF INVENTION: Prokaryotes  
 FILE REFERENCE: ELITRA.011A



```

/ CURRENT APPLICATION NUMBER: US/09/815,242
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 13665
/ LENGTH: 511
/ TYPE: PRT
/ ORGANISM: Salmonella typhi
/ FEATURE:
/ NAME/KEY: VARIANT
/ LOCATION: (1)...(511)
/ OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-13665

```

```

Query Match      64.0%; Score 1783; DB 9; Length 511;
Best Local Similarity 67.4%; Pred. No. 1,5e-159;
Matches 341; Conservative 64; Mismatches 101; Indels 0; Gaps 0;

```

```

QY 2 TKFTITGGVYSSLGKGIASLSLAILEDRLKATTTTKLDYINVDGTMSPFOHGEVF 61
DB 3 TNYIFVTGGVYSSLGKGIASLSLAILEDRLKATTTTKLDYINVDGTMSPFOHGEVF 62
QY 62 TEDGAEITDLDGHTERFLKTTMKNNFTTGOVYEQYLNRKRDYLGATVQYTPHITD 121
DB 63 TEDGAEITDLDGHTERFLKTTMKNNFTTGOVYEQYLNRKRDYLGATVQYTPHITD 122
QY 122 IKRRYVSAEGKDVALLVGGTVGDIESLPLETFIRQMGVELGRDRLFIHLTLVPIKS 181
DB 123 IKRRYVSAEGKDVALLVGGTVGDIESLPLETFIRQMGVELGRDRLFIHLTLVPIKS 182
QY 182 AGELEKPTQSHVVELRTIGIQPDILCRSEQPIASERKRIALFTNVAKAVIATDAD 241
DB 183 AGEVTKPTQSHVVELRTIGIQPDILCRSDRAVPANERAKIALFCNVPKAVISMKDVD 242
QY 242 TIYRIPLLRBQGLDDEVVDQLRLDVPADLSAMEKVVDGLTHTFDEVSAIVGKYVDHT 301
DB 243 SIYKIPGLKQGLDDYICRFPALNCPANLSEVQIYEBANAGSVITIGMVGKILEP 302
QY 302 DAVSINRALTHAGHTRKQVQISYIDSETIEAGTKAKQNDAILVPGFGRGVBECKI 361
DB 303 DAVSIVRALHGHGKKNVYVNIKLIDSDVETRGVAILKDLAILRPGFGRGVBECKI 362
QY 362 STVFARENKIPYIGICIGMOSAVIEFARNVVGEGASTFELPKSPHPTIGLITEMMD 421
DB 363 ATAAVARENKIPYIGICIGMOSAVIEFARNVVGEGASTFELPKSPHPTIGLITEMMD 422
QY 422 AGEVTRDEDSDLGHTERFLKQKCRKLKADSLAPOLYKQDVITERRHRYEFNNQYLK 481
DB 423 DGNVAVSEKSDLDGHTERFLKQKCRKLKADSLAPOLYKQDVITERRHRYEFNNQYLK 482
QY 482 AAGKTFSGKSLDGRLEVEIIELEPHEP 507
DB 483 AAGKTFSGKSLDGRLEVEIIELEPHEP 508

```

```

RESULT 8
US-09-815-242-12372
; Sequence 12372, Application US/09815242
; Patent No. US20020061565A1

```

```

/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ TITLE OF INVENTION: Prokaryotes
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 12372
/ LENGTH: 536
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
US-09-815-242-12372

```

```

Query Match      56.5%; Score 1576; DB 9; Length 536;
Best Local Similarity 55.6%; Pred. No. 6e-140;
Matches 302; Conservative 93; Mismatches 134; Indels 14; Gaps 4;

```

```

QY 1 MTKFTITGGVYSSLGKGIASLSLAILEDRLKATTTTKLDYINVDGTMSPFOHGEVF 60
DB 1 MTKFTITGGVYSSLGKGIASLSLAILEDRLKATTTTKLDYINVDGTMSPFOHGEVF 60
QY 61 VTDGAEITDLDGHTERFLKTTMKNNFTTGOVYEQYLNRKRDYLGATVQYTPHITD 120
DB 61 VTDGAEITDLDGHTERFLKTTMKNNFTTGOVYEQYLNRKRDYLGATVQYTPHITD 120
QY 121 EIKRRYVSAE--GRDVALIEVGGTVGDIESLPLETFIRQMGVELGRDRLFIHLTLVPI 178
DB 121 EIKRRYVSAE--GRDVALIEVGGTVGDIESLPLETFIRQMGVELGRDRLFIHLTLVPI 178
QY 178 IKRAGELKTKTQSHVVELRTIGIQPDILCRSEQPIASERKRIALFTNVAKAVIATDAD 238
DB 179 IKRAGELKTKTQSHVVELRTIGIQPDILCRSEQPIASERKRIALFTNVAKAVIATDAD 238
QY 238 DAVTIRPILRBQGLDDEVVDQLRLDVP--ADLSAMEKVVDGLTHTFDEVSAIVGKY 297
DB 239 DAVTIRPILRBQGLDDEVVDQLRLDVP--ADLSAMEKVVDGLTHTFDEVSAIVGKY 297
QY 297 DADSIYELPDLQSNQNDYIYKQLQANAKETQQLDNRKQDLYVNNLDGKITIGLVGKY 300
DB 241 DADSIYELPDLQSNQNDYIYKQLQANAKETQQLDNRKQDLYVNNLDGKITIGLVGKY 300
QY 300 VDRHDVYSINRALTHAGHTRKQVQISYIDSETIEAGTKAKQNDAILVPGFGRGVBECKI 356
DB 301 VDRHDVYSINRALTHAGHTRKQVQISYIDSETIEAGTKAKQNDAILVPGFGRGVBECKI 356
QY 356 VEGKISTVFARENKIPYIGICIGMOSAVIEFARNVVGEGASTFELPKSPHPTIGLITEMMD 416
DB 357 VEGKISTVFARENKIPYIGICIGMOSAVIEFARNVVGEGASTFELPKSPHPTIGLITEMMD 416
QY 416 SEGKISATIKARENKIPYIGICIGMOSAVIEFARNVVGEGASTFELPKSPHPTIGLITEMMD 420
DB 417 SEGKISATIKARENKIPYIGICIGMOSAVIEFARNVVGEGASTFELPKSPHPTIGLITEMMD 420
QY 420 EMDVAGELVTRDEDSDLGHTERFLKQKCRKLKADSLAPOLYKQDVITERRHRYEFNNQYLK 476
DB 421 EMDVAGELVTRDEDSDLGHTERFLKQKCRKLKADSLAPOLYKQDVITERRHRYEFNNQYLK 476
QY 476 EMDVAGELVTRDEDSDLGHTERFLKQKCRKLKADSLAPOLYKQDVITERRHRYEFNNQYLK 477
DB 477 EMDVAGELVTRDEDSDLGHTERFLKQKCRKLKADSLAPOLYKQDVITERRHRYEFNNQYLK 477

```

```
DB 471 REGULANGVISGTSPPDRLVEMVEIPIINDFFIACQHPPELSRPNRPHPIFKSFIASL 530
QY 537 KEK 539
DB 531 KYQ 533

RESULT 9
US-09-815-242-12978
/ Sequence 12978, Application US/09815242
/ Patent No. US2002061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ PRIOR FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 12978
/ LENGTH: 536
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
US-09-815-242-12978

Query Match
Best Local Similarity 55.5%; Score 1576; DB 9; Length 536;
Matches 302; Conservative 93; Mismatches 134; Indels 14; Gaps 4;

QY 1 MTKPFIITGVVSSISGKGIASLSAILEDRLVTTTKLDPIYINVPDGTMSPPQHGVEF 60
DB 1 MTKPFIITGVVSSISGKGIASLSGRLIKDGLVTTIQKDPYINVPDGTMSPPQHGVEF 60
QY 61 VTDDGAETDLDLGHYERELTKTTKKNFTTGQYVEOVLNRKRGDYLGATVQYIPIHTD 120
DB 61 VTDDGAETDLDLGHYERELTKTTKKNFTTGQYVEOVLNRKRGDYLGATVQYIPIHTD 120
QY 121 BIKRIVYSAS--GKDVALLIEVGTVGDIBSLPFLFTIRQMGVZLGRDRLFIHTLVVPY 178
DB 121 BIKRIVYSAS--GKDVALLIEVGTVGDIBSLPFLFTIRQMGVZLGRDRLFIHTLVVPY 178
QY 179 IKSAGELTKPTQSHVKEKRLPTIGTOPDLICRSQPIPASRRKIALFTVAEKAVISAI 238
DB 179 IKSAGELTKPTQSHVKEKRLPTIGTOPDLICRSQPIPASRRKIALFTVAEKAVISAI 238
QY 239 DADITYRPLLEEGGLDVLVDQRLDVP--AADISAMEKVVDELTPTEVSIAVGYK 297
DB 239 DADITYRPLLEEGGLDVLVDQRLDVP--AADISAMEKVVDELTPTEVSIAVGYK 297
QY 241 DADSLYELPLOSQNMDDIVIKRLQINAKYETQLEDMKQLDDIVNNLQKTIIGLVK 300
DB 241 DADSLYELPLOSQNMDDIVIKRLQINAKYETQLEDMKQLDDIVNNLQKTIIGLVK 300
QY 298 VDRHDVYSLSNBLIHAGIHRKRVQIYIISSETIEAGRAK--LKNVDALIVPGREG 356
DB 298 VDRHDVYSLSNBLIHAGIHRKRVQIYIISSETIEAGRAK--LKNVDALIVPGREG 356
```

```
DB 301 VSLQDAIYLSVESLKHAAIYPPAKDIDIRMISSSVTDENALVYLDVDGILVPGGPFRA 360
QY 357 VEGKISTVPARENKIPYLGICLQGSQAVIEPAAVVGLEGAHSTFPLPKSEHFVIGLIT 416
DB 361 SEKGISALIKYARENVPFGICLQGLATVYSFHNVLGEGAHSAELDPATPYDILILP 420
QY 417 EEMDAGELVTRDBSDIGSTMPLGAOKCRKADSLAFOLYQXQVITEPHERHRYEFPNOY 476
DB 421 EGYD-----LEDIGSTRGLTIPCSIKEGTLQDVYGAKEIERHRHRYEFPNOY 470
QY 477 EKQLEPAQKESGKSLDRGLVEIIELEPHWFLACQHPPEFTSPRNGHALFSGVPA 536
DB 471 REGULANGVISGTSPPDRLVEMVEIPIINDFFIACQHPPELSRPNRPHPIFKSFIASL 530
QY 537 KEK 539
DB 531 KYQ 533

RESULT 10
US-09-815-242-4984
/ Sequence 4984, Application US/09815242
/ Patent No. US2002061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ PRIOR FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4984
/ LENGTH: 536
/ TYPE: PRT
/ ORGANISM: Enterococcus faecalis
US-09-815-242-4984

Query Match
Best Local Similarity 55.3%; Score 1569; DB 9; Length 536;
Matches 301; Conservative 94; Mismatches 134; Indels 14; Gaps 4;

QY 1 MTKPFIITGVVSSISGKGIASLSAILEDRLVTTTKLDPIYINVPDGTMSPPQHGVEF 60
DB 1 MTKPFIITGVVSSISGKGIASLSGRLIKDGLVTTIQKDPYINVPDGTMSPPQHGVEF 60
QY 61 VTDDGAETDLDLGHYERELTKTTKKNFTTGQYVEOVLNRKRGDYLGATVQYIPIHTD 120
DB 61 VTDDGAETDLDLGHYERELTKTTKKNFTTGQYVEOVLNRKRGDYLGATVQYIPIHTD 120
QY 121 BIKRIVYSAS--GKDVALLIEVGTVGDIBSLPFLFTIRQMGVZLGRDRLFIHTLVVPY 178
DB 121 BIKRIVYSAS--GKDVALLIEVGTVGDIBSLPFLFTIRQMGVZLGRDRLFIHTLVVPY 178
```

```

Db      121 EIKKXIMRAAKTADAVITTEVGTVGDIISLPFEALRQKMGSDSNVYIHTTILPY 180
Qy      179 IKSAGELKTPQHSVKELEKRTIGIOPDILICSEQIPASERRKIALFTNVAEKAVISAI 238
Db      181 LKAAGEMKTPQHSHVKELEKRTIGIOPDILICSEQIPASERRKIALFTNVAEKAVISAI 240
Qy      239 DADTYRIPLLREGLDDLVVDQRLVPAADLSAMERKVVDTLTHPTDEVIAIVGKYV 298
Db      241 DVETLYSIPALQANMOQIVCDHLKLPADMTMTALBEKVTMLKKTKIALVGYKV 300
Qy      299 DHTDAKSLNEALIHAGIHTRHKVOISYIDSETIABGTA-KLKXVDAIIVPGSGRGV 357
Db      301 ELPDAIVSYEALKAGPDPDSIDIDWDSQELTAENVAERIGSADGILVPGGGRGI 360
Qy      358 EGIKSTVPARENKIPYIGICIGMQSAVIEFARNVGLGASHTEFLPKSPHPIGLITE 417
Db      361 EGIKSTVPARENDVPFGICIGMQMACVEFGRNVLGEDAGSAETNPDVNNITLMD 420
Qy      418 NMDAGELVTRDESDIGTMTLGAOKRLKADSLAPOLY-QKQVITERHRRHRYEFNNQY 476
Db      421 -----QENINLGGTTLGLGYPCKLKKGTTAAAYGNEDVQRRHRRHRYEFNNKY 470
Qy      477 LKOLEAAGKFFSGKSLDGLVYIIELEPHWFLACQFHPFTSPRNHALFSGFVAAA 536
Db      471 RQLEFENGIVFSGVSPDNRVLVIVIPKQFVACQFHPILISRRNRQRLIKGFVGAAL 530
Qy      537 KHK 539
Db      531 ANK 533

RESULT 11
US-09-815-242-10727
/ Sequence 10727, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815.242
/ PRIOR FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 10727
/ LENGTH: 536
/ TYPE: PRT
/ ORGANISM: Enterococcus faecalis
US-09-815-242-10727
Query Match          56.3%; Score 1569; DB 9; Length 536;

```

```

Best Local Similarity 55.4%; Pred. No. 2.8e-139;
Matches 301; Conservative 94; Mismatches 134; Indels 14; Gaps 4;

Qy      1 MKKFIPTGVSSSGKGLAASLAIEDRLKVTITKLPDIYVNDPQIMSPPQHGWF 60
Db      1 MKKFIPTGVSSSGKGLVAAISLRLKRLKRLKTTIQKPDYINVDGIMSPPYHGWF 60
Qy      61 VTBDGAEITLDIGHYERPLKTTMTKNNFTTGQVYEQVLRNRRKGDYGAIVQVIPHITD 120
Db      61 VTBDGAEITLDIGHYERFIDIMANKYSVTTKISYEVLRKRKKEYGATVQVIPHITN 120
Qy      121 EIKRKYAESAE--GQDVNLIENVGTVGDIISLPFEALRQKMGSDSNVYIHTTILPY 178
Db      121 EIKKXIMRAAKTADAVITTEVGTVGDIISLPFEALRQKMGSDSNVYIHTTILPY 180
Qy      179 IKSAGELKTPQHSVKELEKRTIGIOPDILICSEQIPASERRKIALFTNVAEKAVISAI 238
Db      181 LKAAGEMKTPQHSHVKELEKRTIGIOPDILICSEQIPASERRKIALFTNVAEKAVISAI 240
Qy      239 DADTYRIPLLREGLDDLVVDQRLVPAADLSAMERKVVDTLTHPTDEVIAIVGKYV 298
Db      241 DVETLYSIPALQANMOQIVCDHLKLPADMTMTALBEKVTMLKKTKIALVGYKV 300
Qy      299 DHTDAKSLNEALIHAGIHTRHKVOISYIDSETIABGTA-KLKXVDAIIVPGSGRGV 357
Db      301 ELPDAIVSYEALKAGPDPDSIDIDWDSQELTAENVAERIGSADGILVPGGGRGI 360
Qy      358 EGIKSTVPARENKIPYIGICIGMQSAVIEFARNVGLGASHTEFLPKSPHPIGLITE 417
Db      361 EGIKSTVPARENDVPFGICIGMQMACVEFGRNVLGEDAGSAETNPDVNNITLMD 420
Qy      418 NMDAGELVTRDESDIGTMTLGAOKRLKADSLAPOLY-QKQVITERHRRHRYEFNNQY 476
Db      421 -----QENINLGGTTLGLGYPCKLKKGTTAAAYGNEDVQRRHRRHRYEFNNKY 470
Qy      477 LKOLEAAGKFFSGKSLDGLVYIIELEPHWFLACQFHPFTSPRNHALFSGFVAAA 536
Db      471 RQLEFENGIVFSGVSPDNRVLVIVIPKQFVACQFHPILISRRNRQRLIKGFVGAAL 530
Qy      537 KHK 539
Db      531 ANK 533

RESULT 12
US-09-815-242-13376
/ Sequence 13376, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815.242
/ PRIOR FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931

```

PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 NUMBER OF SEQ ID NOS: 14110  
 SOFTWARE: FASTSEQ for Windows Version 4.0  
 SEQ ID NO 13376  
 LENGTH: 535  
 TYPE: PR  
 ORGANISM: Streptococcus pneumoniae  
 US-09-815-242-13376

Query Match 56.0%; Score 1561; DB 9; Length 535;  
 Best Local Similarity 54.3%; Pred. No. 1,6e-138;  
 Matches 294; Conservative 108; Mismatches 125; Indels 14; Gaps 4;

```

QY 2 TKTFTIGVSSSGKGIASLAILEDRLKVTITKDPYINVDPGTMSPFQHGVEV 61
DB 3 TKYTFVGVSSSGKGIASLAILEDRLKVTITKDPYINVDPGTMSPFQHGVEV 62
QY 62 TEDGAEFDLDGHERFLKTMKNNFTGVQVEQLRNERKADYAGATVQVPHITDE 121
DB 63 TDDGAEFDLDGHERFLDINKVSNVTGKIYSEVLRKEREGETVIGATVQVPHITDA 122
QY 122 IKRNVESA--EGDVALIEVGTVGDIESTPLEFTIRQMGVLEGRDALFTHTLTVPTI 179
DB 123 LKEXIKRAALTDSVITTEVGTVGISLPLEFLRQKADVGADNVMYIHTLPL 182
QY 180 KSAGELKKTPTQSHVSKELRTIGIOPDLICRSQPIPAERKIALFTVAEKAVISAID 239
DB 183 KKAQEMKTKPTQSHVSKELRGLGLOPNLVIRTEBPAGGIKNNLAQFCVAPRAVIESID 242
QY 240 AOTIYRIPLIREGDDLVVDLRDVPADLSAMEKTVDGILHTPTDEVSIALVSKYVD 299
DB 243 VEHLYQIPLNIOAQMOQIVCDHLKIDAPADMTESAMWDKVMNLKQVKSILVGYE 302
QY 300 HTDVKSLNEALIHAGIHTHRKVOISYIDSETTEAGCTAK-LKNVALIVPGGFERGVE 358
DB 303 LQDAIVSVENLKSGVNDVENVKIMVANDVTAENVALLSDAGIIVPGGFERGVE 362
QY 359 GKISTVPARENKIPLYGICGQSAVIEFARNVLEGASTFELPKSPHPVIGLITTEM 418
DB 363 GKIOAIRVARENDDVPMGVCIGQQLTCEIPARHVLGHEGANSALPETYPTIIDMRQ 422
QY 419 MDEAGELVTRDESDLGIMRLGAQKCLKADSLAPQY-QKDVITEHRHREPNQVL 477
DB 423 ID-----IEDWGCTRLGAYPSKLRGSLAAAYNQEVPQRHHRHREPNNAFR 472
QY 478 KQLEBAQMKESGKSLDGLRVLIELPEHPWFLACQFHPFTSTPRNGHALFSGFVZAAK 537
DB 473 EQFLAAGFVSGVSPDRRLVEIYVIEPNKPFVACQYHPBELSSRPNRPESLYTAFTVAIVE 532
QY 538 H 538
DB 533 N 533
  
```

## RESULT 13

US-09-815-242-11320

Sequence 11320, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zvekind, Judith W.

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

TITLE REFERENCE: ELITRA-011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 NUMBER OF SEQ ID NOS: 14110  
 SOFTWARE: FASTSEQ for Windows Version 4.0  
 SEQ ID NO 11320  
 LENGTH: 538  
 TYPE: PR  
 ORGANISM: Helicobacter pylori  
 US-09-815-242-11320

Query Match 51.8%; Score 1444.5; DB 9; Length 538;  
 Best Local Similarity 50.8%; Pred. No. 1.7e-127;  
 Matches 271; Conservative 112; Mismatches 147; Indels 3; Gaps 2;

```

QY 3 KFIPTGVSSSGKGIASLAILEDRLKVTITKDPYINVDPGTMSPFQHGVEVPT 62
DB 5 KFIPTGVSSSGKGISSSIATLLQHCNVQVSIKIDIPYINIDPGTMSPLRHGVPT 64
QY 63 EDGAEFDLDGHERFLKTMKNNFTGVQVEQLRNERKADYAGATVQVPHITDEI 122
DB 65 SDGAEFDLDGHERFLRNLRTLNNTGQISSVTEMRKSGVTKIIVPHITDEI 124
QY 123 KRRVESAEGDVALIEVGTVGDIESTPLEFTIRQMGVLEGRDALFTHTLTVPYIKSA 182
DB 125 KRRIKSAAGDPLIVGVGTVGDMGEMFLKALRQKELGNEKVINAVTLPIYIQT 184
QY 183 GELKTKPTQSHVSKELRTIGIOPDLICRSQPIPAERKIALFTVAEKAVISAIDAJ 242
DB 185 NELTKTKPTQSHVSKELRLGAVTPOILARSKPLDKELKNNLASCVEQDSVLAATDYS 244
QY 243 IYRIPLIREGDDLVVDLRDVPADLSAMEKTVDGILHTPTDEVSIALVSKYVDHT 302
DB 245 IYACPILFLQSGIITPLARRPNLKHPRMAANTLVEKLIAPKHVKIGFVGKYLSE 304
QY 303 AYKSLNEALIHAGIHTHRKVOISYIDSETTEAGCTAKLKNVDAIIVPGGFERGVEGKIS 362
DB 305 SYKSLIENALIHAGIHTDQVNIEMLDESENFNER--TDLBGVDALIVPGGGERGIGKIC 362
QY 363 TVRFARENKIPLYGICGQSAVIEFARNVLEGASTFELPKSPHPVIGLITTEMMDA 422
DB 363 ALOPARLEKLPFLDGLICGQDLAIVERERNVLEKANSSTPRNRCRVPVYLLGDMDQW 422
QY 423 GELVTRDESDLGIMRLGAQKCLKADSLAPQYQKDVITEHRHREPNQVYLKQLEA 482
DB 423 HQQVARTVNSPLGIMRLGEYCEIMNLSLEKAYKCPISIKERHRRHYELNPKRDEWEN 482
QY 483 AGMFGSKSLDGLRVLIELPEHPWFLACQFHPFTSTPRNGHALFSGFVEEA 535
DB 483 KGLNVVVG-SNHLIEALIEHDHFPVGVQFHPFTSRQSPNPIIIDPFKSA 534
  
```

## RESULT 14

US-10-156-761-14031

Sequence 14031, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

```

; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14031
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14031

```

```

Query Match          51.4%; Score 1432.5; DB 15; Length 549;
Best Local Similarity 52.5%; Pred. No. 2.3e-126;
Matches 287; Conservative 92; Mismatches 153; Indels 15; Gaps 7;

```

```

QY 2 TKFIFITGCVVSSLSCKGIAASSLAILEDRLKVTITLDPYINVDPGTMSPFQHGGEV 61
DB 8 TKHIFVTGCVVSSLSCKGLTASLSGLKARGLRVMMQKLDYLVNDPOTMNPFGHGGEV 67
QY 62 TEDGATDLGHTERFLKTTMTKNNFTTGOVVEOVLNRKRDYIGATVQVPHITDE 121
DB 68 TNDGATDLGHTERFLDRDLDSANVTGOVSVYIAERKRGEGYAGDVQVPHITNE 127
QY 122 IKRRVYESAEK-DVALIEVGTVGDIESLPLETIRQMGVELGRDALFIHLTVPIK 180
DB 128 IKHRRBATTEVDVYIEVGTVGDIESLPLETIRQMGVELGRDALFIHLTVPIK 187
QY 181 SAGELKTPQTHSVKELRTTIGIOPDLICRSEOPIPASERKIALFTNVAEKVISAIDA 240
DB 188 PSGLKTPQTHSVKELRTTIGIOPDLICRSEOPIPASERKIALFTNVAEKVISAIDA 247
QY 241 DTIYRIPLLRBOGLDVLVDOLRLDVPADLSAMEKVDGLTPTDEVSIALVGVKVDH 300
DB 248 RSIYDIPKTVGEGDLAVVARKLPPFDVDVMTWDDLDLVKPEHEIMLALVKTIDL 307
QY 301 TDVAKSLNEALIHAGIHTRHVKQISYIDSETIE--ABGTAKLKNVDAIIVPGGGRGV 358
DB 308 PDAYLSVTEALPAGGFARARVAKIKWTSDDCKTPAGAGQGLGVDAICIPGGGRGV 367
QY 359 GKISVRRARENKTPYIGICGMSAVIEFARNVGLGASHTFPLKSPHYIGLITEN 418
DB 368 GKVGAIQYAREHRIPPLGLCLGQCIVIEARNLADIIPDNSTEFDSATGHPVISTMAAQ 427
QY 419 MD-ENGSLVTRDESDSLGTMRLGAOKCRLLKADSLAFQLYO-KDVIETERRHRYERNVOY 476
DB 428 LDIVAGS-----GDMGTMRLGMYPKLAESIVABVDKEVEYERRHRYERNVNY 480
QY 477 LKQLE-AAAGKFSKSLDGLVETIEPE--HPWFLACOPHPEPTSPRGHALFSGFVE 533
DB 481 RAELBKRAGLDFSGTSDGKLVEYVEYPREVHPYLVATQAHPELRSPTPHPLFAGLVK 540
QY 534 AAARKHT 540
DB 541 AAVERTK 547

```

```

RESULT 15
US-09-738-626-5066
; Sequence 5066, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAMA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO

```

```

; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentin ver. 3.0
; SEQ ID NO 5066
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5066

```

```

Query Match          49.2%; Score 1372; DB 10; Length 554;
Best Local Similarity 51.9%; Pred. No. 1.3e-120;
Matches 283; Conservative 78; Mismatches 166; Indels 18; Gaps 6;

```

```

QY 2 TKFIFITGCVVSSLSCKGIAASSLAILEDRLKVTITLDPYINVDPGTMSPFQHGGEV 61
DB 10 TKHIFVTGCVVSSLSCKGLTASLSGLKARGLRVMMQKLDYLVNDPOTMNPFGHGGEV 69
QY 62 TEDGATDLGHTERFLKTTMTKNNFTTGOVVEOVLNRKRDYIGATVQVPHITDE 121
DB 70 TNDGATDLGHTERFLDRDLDSANVTGOVSVYIAERKRGEGYAGDVQVPHITNE 129
QY 122 IKRRVYESAE-----GVDVALIEVGTVGDIESLPLETIRQMGVELGRDALFIHLTV 175
DB 130 IKARLSNGEPDAGNADVDVISEVGTVGDIESLPLETIRQMGVELGRDALFIHLTV 189
QY 176 VPIYSAGELKTPQTHSVKELRTTIGIOPDLICRSEOPIPASERKIALFTNVAEKVISAIDA 235
DB 190 VPIYATSGELKTPQTHSVKELRTTIGIOPDLICRSEOPIPASERKIALFTNVAEKVISAIDA 249
QY 236 SAIDMTIYRIPLLRBOGLDVLVDOLRLDVPADLSAMEKVDGLTPTDEVSIALVGVKVDH 295
DB 250 SCDDSSSTYINPDLVYREHDTPTIIRLGLPFRVDVSTWHDLLERVNNPRLHITVIG 309
QY 296 KYVDHTDAVYKSLNEALIHAGIHTRHVKQISYIDSETIEABGTA--KLKNVDAIIVPGGGRGV 353
DB 310 KYIDLPDVLVSVVAEVAAGVANNTRTNKIKITSDDCETPSGAMKALSGDLAVVPGGGRGV 369
QY 364 ERGVGKISTVRRARENKTPYIGICGMSAVIEFARNVGLGASHTFPLKSPHYIGLITEN 413
DB 370 IRGIBKIGAITPARENKIPPLGLCLGQCIVIEARNLADIIPDNSTEFDSATGHPVISTMAAQ 428
QY 414 LITMMDEAGELVTRDESDSLGTMRLGAOKCRLLKADSLAFQLYO-KDVIETERRHRYERNVOY 473
DB 429 TMEE-----OKAVSGADLGTMRLGAVPATLEBSGLVAFYIGTEVEYERRHRYERNVNY 482
QY 474 NOYLKOL-EAAGKFSKSLDGLVETIEPE--HPWFLACOPHPEPTSPRGHALFSGFVE 530
DB 483 NAYRAQIAESDGLVSGTSDGKLVEYVEYPREVHPYLVATQAHPELRSPTPHPLFAGLVK 542
QY 531 FVEAA 535
DB 543 LVKTA 547

```

```

Search completed: January 29, 2004, 16:21:18
Job time : 38.1346 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: January 29, 2004, 15:54:19 ; Search time 15.7685 Seconds  
(without alignments)  
3044.503 Million cell updates/sec

Title: US-09-941-947a-10

Perfect score: 1199  
Sequence: 1 MNPTICQWAVPAAGVGRKM.....IKTRPBDLALQFYEQQA 231

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched:

789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1199	100.0	231	10	US-09-934-903-6 Sequence 6, Appl1
2	1199	100.0	231	10	US-09-934-868-68 Sequence 68, Appl1
3	1199	100.0	231	11	US-09-941-947a-10 Sequence 10, Appl1
4	555.5	46.3	234	9	US-09-815-242-11931 Sequence 11931, A
5	530.5	44.2	225	9	US-09-815-242-11076 Sequence 11076, A
6	529.5	43.4	254	9	US-09-815-242-11722 Sequence 11722, A
7	520.5	43.4	236	9	US-09-815-242-10273 Sequence 10273, A
8	520.5	43.4	236	10	US-09-792-251-11 Sequence 11, Appl1
9	499.5	41.7	236	9	US-09-815-242-13867 Sequence 13867, A
10	297	24.8	232	10	US-09-792-251-8 Sequence 8, Appl1
11	283.5	23.6	302	10	US-09-935-943-8 Sequence 8, Appl1
12	255.5	21.3	232	12	US-10-128-713A-6 Sequence 6, Appl1
13	245.5	20.5	250	15	US-10-156-761-11503 Sequence 11503, A
14	240.5	20.1	218	12	US-10-289-762-621 Sequence 621, App
15	223.5	18.6	256	10	US-09-738-626-6422 Sequence 6422, App

5	217.5	18.1	238	9	US-09-815-242-12471 Sequence 12471, A
7	214.5	17.9	227	9	US-09-815-242-5711 Sequence 5711, App
8	214.5	17.9	238	9	US-09-876-181-2 Sequence 2, Appl1
9	209	17.4	235	9	US-09-815-242-13256 Sequence 13256, A
0	209	17.4	235	10	US-09-792-251-2 Sequence 2, Appl1
1	198.5	16.6	238	9	US-09-876-182-2 Sequence 5707, App
2	193.5	16.1	236	9	US-09-815-242-5707 Sequence 12467, A
3	193.5	16.1	242	9	US-09-815-242-12467 Sequence 628, App
4	126	10.5	125	9	US-09-867-550-658 Sequence 21638, A
5	105.5	8.8	291	12	US-10-369-493-21638 Sequence 10937, A
6	98.5	8.2	461	9	US-09-815-242-10937 Sequence 5017, App
7	98.5	8.2	783	10	US-09-738-626-5017 Sequence 18528, A
8	98	8.2	327	12	US-10-369-493-18528 Sequence 20306, A
9	94	7.8	327	12	US-10-369-493-20306 Sequence 17444, A
0	93	7.8	297	12	US-10-312-273-245 Sequence 245, App
1	93	7.8	297	12	US-10-289-762-487 Sequence 487, App
2	93	7.8	357	12	US-10-369-493-21686 Sequence 21686, A
3	91	7.6	419	12	US-10-369-493-1153 Sequence 1153, App
4	90.5	7.5	285	12	US-10-369-493-1153 Sequence 10400, A
5	90	7.5	456	9	US-09-815-242-10400 Sequence 264, App
6	89.5	7.5	381	16	US-10-080-170-264 Sequence 21535, A
7	89.5	7.5	385	12	US-10-369-493-21535 Sequence 528, App
8	89	7.4	154	9	US-09-939-960-528 Sequence 8238, App
9	89	7.4	154	15	US-10-128-714-8238 Sequence 13693, A
0	89	7.4	1317	12	US-10-369-493-13693 Sequence 23443, A
1	88.5	7.4	361	12	US-10-369-493-23443 Sequence 4551, App
2	88.5	7.4	485	10	US-09-738-626-4551 Sequence 11824, A
3	88	7.3	394	12	US-10-369-493-11824 Sequence 14528, A
4	87.5	7.3	443	15	US-10-156-761-14528 Sequence 14505, A
5	87.5	7.3	831	15	US-10-156-761-14505

## ALIGNMENTS

REST T 1	US-0-934-903-6	Sequence 6, Application US/09934903
US-0-934-903-6	Sequence 6, Application US/09934903	
Pa	ent No. US20020102690A1	
Ge	eral INFORMATION:	
PLICANT:	Kofias, Mattheos	
PLICANT:	Odum, J. Martin	
PLICANT:	Schenzle, Andreas J.	
PLICANT:	No. US20020102690A1, Kelley C.	
PLICANT:	Tomb, Jean-Francois	
PLICANT:	Rouviere, Pierre	
PLICANT:	Picciaglio, Stephen	
PLICANT:	Cheng, Qiong	
FILE OF INVENTION:	Genes Involved in Isoprenoid Compounds Production	
IE REFERENCE:	CL1646 US NA	
RENT APPLICATION NUMBER:	US/09/934,903	
RENT FILING DATE:	2001-08-22	
FOR APPLICATION NUMBER:	60/229,907	
FOR FILING DATE:	September 1, 2001	
WER OF SEQ ID NOS:	24	
FTNAME:	Microsoft Office 97	
ID NO 6		
ENGT:	231	
YPE:	PRT	
GANISM:	Methylomonas 16a	
ATURE:		
HER INFORMATION:	Amino acid sequences encoded by ORF3	
US-0-934-903-6		
QY	TY Match	100.0%; Score 1199; DB 10; Length 231;
QY	Local Similarity	100.0%; Pred. No. 2.1e-115; Indels 0; Gaps 0;
QY	ches 231; Conservative	
QY	1 MNPTICQWAVPAAGVGRKMADRPKQYITPLAKTIVIRHTITRLLESAPFQVAIVSVE 60	
DB	1 MNPTICQWAVPAAGVGRKMADRPKQYITPLAKTIVIRHTITRLLESAPFQVAIVSVE 60	
QY	61 DRYWELSTAKPDITTAAGKERADSVLSAKLEDTIASENDVTVHDAARPCITGSDI 120	

```

Db      61 DPWPELSIAKHPIITAPGKERADSVLSALKALEDIASNDWVLVHDAAPCLTGSDI 120
Qy      121 HLOIDTLKNDPVGGIILASSHDTLKAVDGTITATIDRGHWRALTPOMFKYGMRLDAQ 180
Db      121 HLOIDTLKNDPVGGIILASSHDTLKAVDGTITATIDRGHWRALTPOMFKYGMRLDAQ 180
Qy      181 RTEGNPAVTDBASALELGHKPKIVGGRPNIKITRPEDLALAQFYMEQQA 231
Db      181 RTEGNPAVTDBASALELGHKPKIVGGRPNIKITRPEDLALAQFYMEQQA 231

```

```

RESULT 2
US-09-934-868-68
; Sequence 68: Application US/09941947A
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: C11936 US NA
; CURRENT APPLICATION NUMBER: US/09/934, 868
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229, 858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 68
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Methylobionas 16a
; FEATURE:
; OTHER INFORMATION: Amino acid sequences encoded by ISPD
US-09-934-868-68

```

```

Query Match      100.0%; Score 1199; DB 10; Length 231;
Best Local Similarity 100.0%; Pred. No. 2.1e-115;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MNPTIQCAVAVPAAGVGRKQADRPQYLPAGKTVIEHTITRLLESDFOKVAVAISVE 60
Db      1 MNPTIQCAVAVPAAGVGRKQADRPQYLPAGKTVIEHTITRLLESDFOKVAVAISVE 60
Qy      61 DPWPELSIAKHPIITAPGKERADSVLSALKALEDIASNDWVLVHDAAPCLTGSDI 120
Db      61 DPWPELSIAKHPIITAPGKERADSVLSALKALEDIASNDWVLVHDAAPCLTGSDI 120
Qy      121 HLOIDTLKNDPVGGIILASSHDTLKAVDGTITATIDRGHWRALTPOMFKYGMRLDAQ 180
Db      121 HLOIDTLKNDPVGGIILASSHDTLKAVDGTITATIDRGHWRALTPOMFKYGMRLDAQ 180
Qy      181 RTEGNPAVTDBASALELGHKPKIVGGRPNIKITRPEDLALAQFYMEQQA 231
Db      181 RTEGNPAVTDBASALELGHKPKIVGGRPNIKITRPEDLALAQFYMEQQA 231

```

```

RESULT 3
US-09-941-947a-10
; Sequence 10: Application US/09941947A
; Publication No. US20030003528A1
; GENERAL INFORMATION:
; APPLICANT: Bryzotowicz, Patricia C.
; APPLICANT: Cheng, Qiong
; APPLICANT: DiCosimo, Deana J.
; APPLICANT: Koffas, Mattheos
; APPLICANT: Miller, Edward S. Jr.
; APPLICANT: Odom, J. Martin
; APPLICANT: Picataggio, Steve
; APPLICANT: Rouviere, Pierre E.
; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
; FILE REFERENCE: C11903 US NA
; CURRENT APPLICATION NUMBER: US/09/941, 947A

```

```

; CURRENT FILING DATE: 2001-09-01
; RIOR APPLICATION NUMBER: 60/229, 907
; RIOR FILING DATE: 2000-09-01
; RIOR APPLICATION NUMBER: 60/229, 858
; RIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 10
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Methylobionas 16a
US-9-941-947a-10

```

```

Query Match      100.0%; Score 1199; DB 11; Length 231;
Best Local Similarity 100.0%; Pred. No. 2.1e-115;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MNPTIQCAVAVPAAGVGRKQADRPQYLPAGKTVIEHTITRLLESDFOKVAVAISVE 60
Db      1 MNPTIQCAVAVPAAGVGRKQADRPQYLPAGKTVIEHTITRLLESDFOKVAVAISVE 60
Qy      61 DPWPELSIAKHPIITAPGKERADSVLSALKALEDIASNDWVLVHDAAPCLTGSDI 120
Db      61 DPWPELSIAKHPIITAPGKERADSVLSALKALEDIASNDWVLVHDAAPCLTGSDI 120
Qy      121 HLOIDTLKNDPVGGIILASSHDTLKAVDGTITATIDRGHWRALTPOMFKYGMRLDAQ 180
Db      121 HLOIDTLKNDPVGGIILASSHDTLKAVDGTITATIDRGHWRALTPOMFKYGMRLDAQ 180
Qy      181 RTEGNPAVTDBASALELGHKPKIVGGRPNIKITRPEDLALAQFYMEQQA 231
Db      181 RTEGNPAVTDBASALELGHKPKIVGGRPNIKITRPEDLALAQFYMEQQA 231

```

```

REF ID 4
US-9-815-242-11931
; Sequence 11931, Application US/09845242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zysek, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trivick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE REFERENCES: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815, 242
; CURRENT FILING DATE: 2001-03-21
; RIOR APPLICATION NUMBER: 60/191, 078
; RIOR FILING DATE: 2000-03-21
; RIOR APPLICATION NUMBER: 60/206, 848
; RIOR FILING DATE: 2000-05-23
; RIOR APPLICATION NUMBER: 60/207, 727
; RIOR FILING DATE: 2000-05-28
; RIOR APPLICATION NUMBER: 60/242, 578
; RIOR FILING DATE: 2000-10-23
; RIOR APPLICATION NUMBER: 60/253, 625
; RIOR FILING DATE: 2000-11-27
; RIOR APPLICATION NUMBER: 60/257, 931
; RIOR FILING DATE: 2000-12-22
; RIOR APPLICATION NUMBER: 60/269, 308
; RIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO: 11931
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-9-815-242-11931

```

Query Match 46.3%; Score 555.5; DB 9; Length 234;  
 Best Local Similarity 54.1%; Pred. No. 5.3e-49;  
 Matches 119; Conservative 26; Mismatches 72; Indels 3; Gaps 3;

QY 8 WAWPAAGVGRKQADRPQYPLAGKTVIEHTLTRLESDFQVAVASVDEPYWPEL 67  
 10 WTVIPAGVGRKQADRPQYPLAGKTVIEHTLTRLESDFQVAVASVDEPYWPEL 69

DB 68 SIAPKPDITAPGKERADSVLSA-LKALEDIASNDVTVHDAAPCLTGSNHIQIDTL 126  
 70 DCAASRHOBAAGKERAAGSVLNGILRLIEGAQADWVTVHDAARPNLTRGDIDRLIEE 129

QY 127 LKNDVGGILASSDHTLKHYDG-TITATIDRKHYVRLTPQWEXYGMALDQ-RTGEG 184  
 130 LAEDVGGILAVPADITLKSDRGRVSEITDRSVYMLATPQWFRGLALRALDLVA 189

DB 185 NPATVDEASALELGHKPRIVEGRPDNIKITRPEDLALAQ 224  
 190 GVAITDEASAMWAGYAPKLVGGRADNLKITPDELRLQ 229

## RESULT 5

US-09-815-242-11076  
 Sequence 11076, Application US/09815242  
 Patent No. US20020061569A1  
 GENERAL INFORMATION:  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari L.  
 APPLICANT: Zyskind, Judith W.  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John D.  
 APPLICANT: Carr, Grant J.  
 APPLICANT: Yamamoto, Robert T.  
 APPLICANT: Xu, H. Howard  
 TITLE OF INVENTION: Identification of Essential Genes in  
 TITLE OF INVENTION: Prokaryotes  
 FILE REFERENCE: ELITRA.011A  
 CURRENT APPLICATION NUMBER: US/09/815.242  
 CURRENT FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 NUMBER OF SEQ ID NOS: 14110  
 SOFTWARE: FastSeq for windows Version 4.0  
 SEQ ID NO 11076  
 LENGTH: 225  
 TYPE: PRT  
 ORGANISM: Haemophilus influenzae

Query Match 44.2%; Score 530.5; DB 9; Length 225;  
 Best Local Similarity 52.5%; Pred. No. 1.9e-46;  
 Matches 117; Conservative 28; Mismatches 65; Indels 13; Gaps 6;

QY 9 AVPAAGVGRKQADRPQYPLAGKTVIEHTLTRLESDFQVAVASVDEPYWPEL 68  
 7 AVPAAGVGRKQADRPQYPLAGKTVIEHTLTRLESDFQVAVASVDEPYWPEL 66

DB 69 IAKHPDITAPGKERADSVLSA-LKALEDIASNDVTVHDAAPCLTGSNHIQIDTL 128  
 67 L-DEKIDLVGGITRAESVINGLNA---IAEKNAWLVHDAARPCIQ---HADIDKXL 117

QY 129 --NDVGGILASSDHTLKHYDG-TITATIDRKHYVRLTPQWEXYGMALDQ-RTGEG 184  
 118 AIEDQKQILALPYDTITKRADNQCIVKTEDRSQWQAMTPQWFRGLALRALDLVA 177

DB 185 NPATVDEASALELGHKPRIVEGRPDNIKITRPEDLALAQ 227  
 178 GANITDEASALELGHKPRIVEGRPDNIKITRPEDLALAQ 220

## RESULT 6

US-09-815-242-11722  
 Sequence 11722, Application US/09815242  
 Patent No. US20020061569A1  
 GENERAL INFORMATION:  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari L.  
 APPLICANT: Zyskind, Judith W.  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John D.  
 APPLICANT: Carr, Grant J.  
 APPLICANT: Yamamoto, Robert T.  
 APPLICANT: Xu, H. Howard  
 TITLE OF INVENTION: Identification of Essential Genes in  
 TITLE OF INVENTION: Prokaryotes  
 FILE REFERENCE: ELITRA.011A  
 CURRENT APPLICATION NUMBER: US/09/815.242  
 CURRENT FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 NUMBER OF SEQ ID NOS: 14110  
 SOFTWARE: FastSeq for windows Version 4.0  
 SEQ ID NO 11722  
 LENGTH: 254  
 TYPE: PRT  
 ORGANISM: Klebsiella pneumoniae

Query Match 44.2%; Score 529.5; DB 9; Length 254;  
 Best Local Similarity 51.1%; Pred. No. 2.9e-46;  
 Matches 114; Conservative 30; Mismatches 70; Indels 9; Gaps 5;

QY 9 AVPAAGVGRKQADRPQYPLAGKTVIEHTLTRLESDFQVAVASVDEPYWPEL 68  
 33 AVPAAGVGRKQADRPQYPLAGKTVIEHTLTRLESDFQVAVASVDEPYWPEL 92

DB 69 IAKHPDITAPGKERADSVLSA-LKALEDIASNDVTVHDAAPCLTGSNHIQIDTL 128  
 93 IAKHPDITAPGKERADSVLSA-LKALEDIASNDVTVHDAAPCLTGSNHIQIDTL 148

QY 129 NPVGGILASSDHTLKHYDG-TITATIDRKHYVRLTPQWEXYGMALDQ-RTGEG 184  
 149 TSVGGILAVPADITLKSDRGRVSEITDRSVYMLATPQWFRGLALRALDLVA 208

DB 185 NPATVDEASALELGHKPRIVEGRPDNIKITRPEDLALAQ 227  
 209 ATITDEASALELGHKPRIVEGRPDNIKITRPEDLALAQ 250

## RESULT 7

US-09-815-242-10273



```

Sequence 10273, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haeselbeck, Robert
APPLICANT: Ohlsen, Kari E.
APPLICANT: Zykkind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10273
LENGTH: 236
TYPE: PRI
ORGANISM: Escherichia coli
US-09-815-242-10273

Query Match      43.4%; Score 520.5, DB 9; Length 236;
Best Local Similarity 49.8%; Pred. No. 2,2e-45;
Matches 111; Conservative 31; Mismatches 72; Indels 9; Gaps 4;

QY 9 AVVPAAGGCKMNOADRPKQYVLPAGKTVIEHTLTLLSDAFQKVAVAISVDPYMPKLS 68
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 10 AVVPAAGGRRKQQTCPKQYLSIGNQTILSHSVHALLHPRVKRVVIALSPDSRPAQLP 69
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 69 IAKHPDITAPGKGRADSVLSALKALEDIASENMDVTVHDAAPCLTGSDDHIDITLK 128
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 70 LANRQITVVOGGRADSVLGLKA---AGDAQWVLVHDAARPCIHODDARLLALUSR 125
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 129 NDPVGGILALSHDPLKND--GDITITNIDKHWBALTPQWFKKGMRLAQQR--TEG 184
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 126 TSRTGIIAAPVPRDTMKKAEPEKNALIAHTVDNRNGHMLHPFPFRELHDLCLTRALNG 185
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 185 NPATVDEASALELGHKKPIVEGRPDNIKITRPEDIALAQFYM 227
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 186 -ATITDEASALEYCGETHPOLVWGRADNMIKVTPEPDLAABFYL 227
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

RESULT 8
US-09-792-251-11
; Sequence 11, Application US/09792251
; Patent No. US20020160364A1
GENERAL INFORMATION:
APPLICANT: Fritz, Christian
APPLICANT: Youngman, Philip
APPLICANT: Guzman, Luz-Maria
TITLE OF INVENTION: USE OF YACM AND YOCJ, ESSENTIAL BACTERIAL GENES AND POLYPEPTIDES
FILE REFERENCE: 06286-140001
CURRENT APPLICATION NUMBER: US/09/792,251
CURRENT FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 28

```

```

SOFTWARE: FastSeq for Windows Version 4.0
; C ID NO 11
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Escherichia coli
us: 9-792-251-11

Query Match          43.4%; Score 520.5; DB 10; Length 236;
Hist Local Similarity 49.8%; Pred. No. 2.2e-45;
Matches 111; Conservative 31; Mismatches 72; Indels 9; Gaps 4

QY 9 AVPPAAGVGRKQADPRKQYIPLAKTYIHTLRLLESAPQKVAIVSEDPYVEIS 68
Db 10 AVPPAAGPRERKQTCSPKQYISIGNQITLESVYALLAEPRVKEVVAIASEPGDSRAQLP 69
QY 69 IAHPPAIIAPGKGRADSVLSAKAEEDIASENDWLVVHDAAPCLGSDIHLQIDTLK 128
Db 70 IAHNPPIYVDSGDERADSVLSAKLA-----AGDAQWLVVHDAAPCLHQQDILARLLALSE 125
QY 129 NDEVGIIALSSHTLTAKVD--GDTITKATIDRKAVWALTPQKFKYGMPLALQR--TEG 184
Db 126 TSRGTGIIAAPEDTIDMKRAEPGRKAIATVDRNGELMHALTPQFPPELLHDCYLRALNEG 185
QY 185 NPAYTDEASATELTGHKRPIVEGRDNIKITRPEDLALAAQFM 227
Db 186 -ATITDEASALEYCGFHQVLVBGRADNINKYIRPEDLALAEFTL 227

```

```

US: 9-815-242-13867
SEQUENCE 13867, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Onlesen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Protekocytes
PRIORITY REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIORITY APPLICATION NUMBER: 60/191,078
PRIORITY FILING DATE: 2000-03-21
PRIORITY APPLICATION NUMBER: 60/206,848
PRIORITY FILING DATE: 2000-05-23
PRIORITY APPLICATION NUMBER: 60/207,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: 60/242,578
PRIORITY FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/269,308
PRIORITY FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 1410
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 13867
LENGTH: 236
TYPE: PRT
ORGANISM: Salmonella typhi
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(236)
OTHER INFORMATION: Xaa = Any Amino Acid
US: 9-815-242-13867

```

Best Local Similarity 48.4%; Pred. No. 3.2e-43;  
 Matches 108; Conservative 32; Mismatches 74; Indels 9; Gaps 4;

QY 9 AVPAAGVGMQADRPQYPLAGKTVIEHTLRLLESDFQKVAVAISVEDPYPELS 68  
 DB 10 AVPAAGVGMQADRPQYPLAGKTVIEHTLRLLESDFQKVAVAISVEDPYPELS 69  
 QY 69 IAKRPDITTAGGERADSVLSALKALEDIASENDWVLYHDAARPCLTGSDIHLQDITLK 128  
 DB 70 IAKRPDITTAGGERADSVLSALKALEDIASENDWVLYHDAARPCLTGSDIHLQDITLK 125  
 QY 129 NDVPVGLIALSHDLTKVD--GDTITATIDRKHWRAALTPQWFKYGMELDALQR--TEG 184  
 DB 126 NSRVGGLIASPVRTMRKGPCKNAIAHYERADLMALTPQFPRELLMDCLTRALNEG 185  
 QY 185 NPATVDEASALELGHKPKIVEGRPNKIKITREDDALAQFYM 227  
 DB 186 -ATTDEASALEYCGFHPLVEGRADNIKVTRFDLALAEFYL 227

## RESULT 10

US-09-792-251-8  
 ; Sequence 8, Application US/09792251  
 ; Patent No. US20020160364A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fritz, Christian  
 ; APPLICANT: Youngman, Philip  
 ; APPLICANT: Guzman, Luz-Maria  
 ; TITLE OF INVENTION: USE OF YACM AND YOEJ, ESSENTIAL BACTERIAL GENES AND POLYPEPTIDES  
 ; TITLE OF INVENTION: AND THEIR USE  
 ; FILE REFERENCE: 06286-140001  
 ; CURRENT APPLICATION NUMBER: US/09/792.251  
 ; CURRENT FILING DATE: 2001-02-23  
 ; NUMBER OF SEQ ID NOS: 28  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 8  
 ; LENGTH: 232  
 ; TYPE: PRT  
 ; ORGANISM: Bacillus subtilis  
 ; US-09-792-251-8

Query Match 24.8%; Score 297; DB 10; Length 232;  
 Best Local Similarity 36.8%; Pred. No. 2.5e-22;  
 Matches 85; Conservative 40; Mismatches 88; Indels 18; Gaps 9;

QY 10 VVPAAGVGMQADRPQYPLAGKTVIEHTLRLLESDFQKVAVAISVEDPYPELS 68  
 DB 6 VVPAAGVGMQADRPQYPLAGKTVIEHTLRLLESDFQKVAVAISVEDPYPELS 64  
 QY 69 IAKRPDITTAGGERADSVLSALKALEDIASENDWVLYHDAARPCLTGSDIHLQDITLK 125  
 DB 65 IAKRPDITTAGGERADSVLSALKALEDIASENDWVLYHDAARPCLTGSDIHLQDITLK 116  
 QY 126 TL--KNDPVG-GIALSHDLTKVDGDTITATIDRKHWRAALTPQWFKYGMELDALQR 182  
 DB 117 ELIASEGTGALIAVAVKQITKRVQDIQVSEITERSLVAVQTPQAFRLSLMKHAA 176  
 QY 183 EGNPAV-TDEASALELGHKPKIVEGRPNKIKITREDDALAQFYM 231  
 DB 177 ERKGLGTGDSALVEOMEGSVRVESGTYNKLTPPDLTSAEAMISES 227

## RESULT 11

US-09-935-943-8  
 ; Sequence 8, Application US/09935943  
 ; Patent No. US20020120963A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Levin, Joshua Z.  
 ; APPLICANT: Megrich, Lynette M.  
 ; APPLICANT: Budziszewski, Gregory J.  
 ; TITLE OF INVENTION: Herbicide Target Genes and Methods  
 ; FILE REFERENCE: PB/5-31509A  
 ; CURRENT APPLICATION NUMBER: US/09/935.943

; CURRENT FILING DATE: 2001-08-23  
 ; FOR APPLICATION NUMBER: PCT/EP01/08910  
 ; FOR FILING DATE: 2001-08-01  
 ; FOR APPLICATION NUMBER: US 60/222,779  
 ; FOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 302  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana  
 ; US-09-935-943-8

QY Match 23.6%; Score 283.5; DB 10; Length 302;  
 Best Local Similarity 34.2%; Pred. No. 9e-21;  
 Matches 80; Conservative 46; Mismatches 83; Indels 25; Gaps 10;

QY 10 VVPAAGVGMQADRPQYPLAGKTVIEHTLRLLESDFQKVAVAISVEDPYPELS 63  
 DB 82 ILLAGGQGRMKMMPKQYIPLLQPIALVSFFPFSRMP-----VKBIVVCCDPFRD 135  
 QY 64 -WPELSIAKHPDITTAGGERADSVLSALKALEDIASENDWVLYHDAARPCLTGSDIHL 122  
 DB 136 IFEYTESIDVLPALPKERQDSVYSGLOEI-DVNSE--LYCINHSARFLVNTDEVY- 191  
 QY 123 QIDTLKNDPVG--IALSHDLTKVDGDTITATIDRKHWRAALTPQWFKYGMELDAL 178  
 DB 192 --KVLKQSAVGAIVLGPAAKATIKVNSDSLVTKTIDRKTEMQPVYKPELLKGF 249  
 QY 179 -LQRTKGNPATVDEASALELGHKPKIVEGRPNKIKITREDDALAQFYM 221  
 DB 250 ELVASEG-LEVTDVDSIVELYKHPVYVQSGSYTNKIKVTPDDLLMERILSEDS 302

## RESULT 12

US-1-128-713A-6  
 ; Sequence 6, Application US/10128713A  
 ; Patent No. US20030170847A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bramucci, Michael G  
 ; TITLE OF INVENTION: Genes Involved in Isoprenoid Compound Production  
 ; FILE REFERENCE: CL-1788  
 ; CURRENT APPLICATION NUMBER: US/10/128,713A  
 ; CURRENT FILING DATE: 2002-04-22  
 ; NUMBER OF SEQ ID NOS: 36  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 6  
 ; LENGTH: 232  
 ; TYPE: PRT  
 ; ORGANISM: Rhodococcus erythropolis  
 ; US-1-128-713A-6

QY Match 21.3%; Score 255.5; DB 12; Length 232;  
 Best Local Similarity 35.0%; Pred. No. 4.8e-18;  
 Matches 78; Conservative 30; Mismatches 96; Indels 19; Gaps 5;

QY 9 AVPAAGVGMQADRPQYPLAGKTVIEHTLRLLESDFQKVAVAISVEDPYPELS 67  
 DB 5 AVPAAGVGMQADRPQYPLAGKTVIEHTLRLLESDFQKVAVAISVEDPYPELS 58  
 QY 68 -----SIKRPDITTAGGERADSVLSALKALEDIASENDWVLYHDAARPCLTGSDIHL 122  
 DB 59 EGVVADLGRASDVVVGGAERTDSVRAGLSA---AGDADFVLYHDAARPCLTGSDIHL 114  
 QY 123 QIDTLKNDPVG-GIALSHDLTKVDGDTITATIDRKHWRAALTPQWFKYGMELDALQR 181  
 DB 115 VVDAIRAGSSAVLPLVPTITIKSVDLGAVTGIPLASERAAVQTPQGSTVDEKSAVD- 173  
 QY 182 TEGNPAVDEASALELGHKPKIVEGRPNKIKITREDDALAQFYM 224  
 DB 174 -AGVNAATDADALVERGVSVQITPPDADALAFKITTLDVLAR 215

## RESULT 13

US-10-156-761-11503  
 ; Sequence 11503, Application US/10156761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OKURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIGAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHIYUKI  
 ; APPLICANT: KATORI, MASAHIRA  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-262  
 ; CURRENT APPLICATION NUMBER: US/10/156,761  
 ; CURRENT FILING DATE: 2002-05-29  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-272697  
 ; PRIOR FILING DATE: 2001-08-02  
 ; NUMBER OF SEQ ID NOS: 15109  
 ; SEQ ID NO 11503  
 ; LENGTH: 250  
 ; TYPE: PR1  
 ; ORGANISM: Streptomyces avermitilis  
 ; US-10-156-761-11503

Query Match 20.5%; Score 245.5; DB 15; Length 250;  
 Best Local Similarity 32.1%; Pred. No. 5,7e-17;  
 Matches 75; Conservative 34; Mismatches 96; Indels 29; Gaps 6;

QY 9 AVVPAAGVGMQADPRKQYLPAGKTVIEHTLRLLESAPQKVAAL-----SV 59  
 DB 16 AVVPAAGVGMQADPRKQYLPAGKTVIEHTLRLLESAPQKVAAL-----SV 75  
 QY 60 EDVPEPLSIKHPDITAPGKERADSVLSALKALDIASENDWLVHDAARPCITGSD 119  
 DB 76 LDAN-----ALPERIDFVVVGSGRSQSVLGLDAL---PPGIDIVLVHDAARPLVPVDT 128  
 QY 120 IHLQIDTLKNDPVGGIILSSHTLKHV-----DGDITITIDRKVWRALTPQMFXY 172  
 DB 129 VDAVIEAVRDGAPVAVPALPLADTVKQVEPAVGPGEVPAVPEPARILRAVOTPO---- 184  
 QY 173 GMLRDALQRTGNGA--VTDEASALELGHKPKIVEGRPNIKITREPDIALAQ 224  
 DB 185 GFDDTLVRAHEITVTDVTDASVWEQDGAHVAVVPCHEBAFVTFPLDIVLAE 238

RESULT 14  
 US-10-289-762-621  
 ; Sequence 621, Application US/10289762  
 ; Publication No. US20040006218A1

; GENERAL INFORMATION:  
 ; APPLICANT: Griffiths, R.  
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
 ; FILE REFERENCE: 9710-003-999  
 ; CURRENT APPLICATION NUMBER: US/10/289,762  
 ; CURRENT FILING DATE: 2003-03-27  
 ; NUMBER OF SEQ ID NOS: 6849  
 ; SEQ ID NO 621  
 ; LENGTH: 218  
 ; TYPE: PR1  
 ; ORGANISM: Chlamydia pneumoniae  
 ; US-10-289-762-621

Query Match 20.1%; Score 240.5; DB 12; Length 218;  
 Best Local Similarity 32.9%; Pred. No. 1.5e-16;  
 Matches 74; Conservative 41; Mismatches 91; Indels 19; Gaps 9;

QY 3 PTIQCNAVPAAGVGMQADPRKQYLPAGKTVIEHTLRLLESAPQKVAALSVEDP 62

DB 7 PMKSLILLSGGQRTGSKLPKQYLPANGTPLYHSL-KILSS--LPQIAEVIWVCDP 63  
 QY 63 YPEPLSIKHPDITAPGKERADSVLSALKALDIASENDWLVHDAARPCITGSD 122  
 DB 64 SYOE-TFOEYVSPALP-GERRQDSVFGLOOV-----SYFVYIHDARRPPIYDETH- 115  
 QY 123 QIDTLKNDPVGGIILSS--HDTLGHVGDITITATIDRKVWRALTPQMFXYGMLRALQ 180  
 DB 116 --DLFTAKIGATLASPIPTIKQNP---VRTLDNDLAIHTPQCIKTEILREGLA 170  
 QY 181 -RTGPAVTDASALELGHKPKIVEGRPNIKITREPDIALAQ 224  
 DB 171 LAKKQTLTVDDIEMAEIIGKPSQVFNHPOIKISTYEDLTIAQ 215

RESULT 15  
 US-9-738-626-6422  
 ; Sequence 6422, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAOKO  
 ; APPLICANT: SENO, AKIHIRO  
 ; APPLICANT: IKEDA, MASARU  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; OPTMARK: Patent in ver. 3.0  
 ; SEQ ID NO 6422  
 ; LENGTH: 256  
 ; TYPE: PR1  
 ; ORGANISM: Corynebacterium glutamicum  
 ; US-9-738-626-6422

Query Match 18.6%; Score 223.5; DB 10; Length 256;  
 Best Local Similarity 30.7%; Pred. No. 1.1e-14;  
 Matches 71; Conservative 43; Mismatches 102; Indels 15; Gaps 7;

QY 9 AVVPAAGVGMQADPRKQYLPAGKTVIEHTLRLLESAPQKVAALSVEDP 67  
 DB 10 ALVLAAGGRTLAGPRTKAFVTLERTLERSLOAMLTSSVBEIILVSPDMEYARDL 69  
 QY 68 ----SIAGRP---ITAPGKERADSVLSALKALDIASENDWLVHDAARPCITGSD 119  
 DB 70 LKRGSLNDPVGAVRVLVHGGGERADSVMAQLAISDQAPPAIVLHDSARALTFPGM 129  
 QY 120 IHLQIDTLKNDPVGGIILSSHTLKHV-----DGDITITIDRKVWRALTPQMFXYGMLRALD 177  
 DB 130 IARVVRKVEGATVAVIPVLPVSDIKKVSPPGQVAVVTPFRAELRAVOTPOGFIISLVA 189  
 QY 178 ALQR--TSGNPAV--VTDEASALELGHKPKIVEGRPNIKITREPDIALAQ 224  
 DB 190 ANEKFPADPNPFGPIPTDASLMEWYGVADVCGDPAFAKVTTEIDMMALAQ 240

Search completed: January 29, 2004, 16:21:16  
 Job time: 16.7685 secs

Thu Jan 29 17:41:17. 2004

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 15:40:33 ; Search time 21.153 Seconds  
(without alignment)  
1736.455 Million cell updates/sec

Title: US-09-941-947A-10

Perfect score: 1199  
Sequence: 1 MNPTICQMAVPAAGVGRK.....IKITRPEDLAAQFVMEQQA 231

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : A\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/emb1/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/emb1/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/emb1/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/emb1/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/emb1/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/emb1/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/emb1/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/emb1/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/emb1/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/emb1/AA1990.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/emb1/AA1991.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/emb1/AA1992.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/emb1/AA1993.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/emb1/AA1994.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/emb1/AA1995.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/emb1/AA1996.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/emb1/AA1997.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/emb1/AA1998.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/emb1/AA1999.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/emb1/AA2000.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/emb1/AA2001.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/emb1/AA2002.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/emb1/AA2003.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1199	100.0	231	23	ABG61584
2	1199	100.0	231	23	AAE22303
3	1199	100.0	231	23	AAU80327
4	555.5	46.3	234	22	AAU36338
5	530.5	44.2	225	22	AAU35483
6	529.5	44.2	254	22	AAU16129
7	524.5	43.7	236	21	AAU47964
8	520.5	43.4	236	21	AAU13772
9	520.5	43.4	236	22	AAU34680

1	520.5	43.4	236	24	AAE31678	Escherichia coli Y
1	499.5	41.7	236	22	AAU38274	Salmonella typhi C
1	411.5	34.3	229	21	AAU74694	Neisseria meningit
1	408.5	34.1	851	22	ABG16344	Novel human diagno
1	397.5	33.2	229	21	AAU74692	Neisseria gonorrhe
1	397.5	33.2	229	24	ABP79613	N. gonorrhoeae am1
1	297	24.8	232	24	AAE31677	Bacillus subtilis
1	283.5	23.6	227	22	AAE70484	Arabisopsis thalia
1	283.5	23.6	263	21	AAU31950	Arabisopsis thalia
1	283.5	23.6	276	21	AAU47685	Arabisopsis thalia
2	283.5	23.6	300	21	AAU311251	Arabisopsis thalia
2	283.5	23.6	302	21	AAU311250	Arabisopsis thalia
2	283.5	23.6	302	22	AAU70485	Arabisopsis thalia
2	283.5	23.6	302	22	AAU39878	Arabisopsis thalia
2	283.5	23.6	302	23	AAU39992	Herbicidally activ
2	283.5	23.6	322	21	AAU47684	Arabisopsis thalia
2	283.5	23.6	324	21	AAU47683	Arabisopsis thalia
2	281.5	23.5	940	22	ABG30158	Novel human diagno
2	270	22.5	232	23	ABG48960	Listeria monocytog
2	258.5	21.6	211	21	AAU311252	Arabisopsis thalia
3	258.5	21.6	211	21	AAU47685	Arabisopsis thalia
3	255.5	21.5	232	24	AAE31684	Rhodococcus erythr
3	254.5	21.2	207	21	AAU47697	Arabisopsis thalia
3	254.5	21.2	209	21	AAU31962	Arabisopsis thalia
3	254.5	21.2	209	21	AAU47696	Arabisopsis thalia
3	249	20.8	152	21	AAU47693	Neisseria meningit
3	246	20.5	236	23	ABG48561	Listeria monocytog
3	243	20.3	298	22	AAU50683	Propionibacterium
3	240.5	20.1	218	20	AAU35202	Chlamydia pneumoni
4	229.5	19.1	239	23	ABP30406	Streptococcus poly
4	229.5	19.1	247	23	ABP27297	Streptococcus poly
4	224.5	18.7	291	23	ABP65462	Bifidobacterium 10
4	223.5	18.6	256	22	AAU32668	C glutamicum prote
4	223.5	18.6	256	22	AAU79384	Corynebacterium gl
4	223.5	18.6	256	22	AAU79461	Corynebacterium gl

## ALIGNMENTS

RESULT 1	ABG6134	3661584 standard; Protein: 231 AA.
ID	3661584;	
XX	7-AUG-2002 (first entry)	
DE	High growth methanotrophic bacterial strain polypeptide #34.	
XX	High growth methanotrophic bacterial strain; C1 carbon substrate; enzyme;	
XX	ethane; methanol; Embden-Meyerhof carbon flux pathway; 16S RNA;	
KW	/phosphatase dependent phosphofructokinase; nitrogen-containing compound;	
KW	amonia; nitrate; nitrite; nitrogen; pigment; oxygen; landfill;	
KW	ethane-containing environment; waste water treatment system; isoprenoid;	
KW	nitrous oxide; terpenoid; animal feed; carotenoid; exopolysaccharide.	
OS	Methylobionas 16a.	
XX		
PN	200220728-AA2.	
XX		
PD	MAR-2002.	
XX		
PF	ADG-2001; 2001MO-US26827.	
XX		
PR	1-SEP-2000; 2000US-229858P.	
XX		
PA	UPO ) DU PONT DE MEMOURS & CO E. I.	
XX		
PI	Offas M, Odum JM, Schenkele A;	
XX		
DR	1-PI; 2002-452200/48.	

DR N-PSDB; ABK83263.

XX New high growth methanotrophic bacterial strain, useful for producing  
PT single cell proteins, grows on a C1 carbon substrate, and comprises a  
PT functional gene encoding in Embden-Meyerhof carbon pathway  
XX  
XX  
PS Claim 11: Page 143-144; 157pp; English.

XX The invention relates to a high growth methanotrophic bacterial strain,  
CC which grows on a C1 carbon substrate e.g. methane and methanol, and  
CC comprises a functional Embden-Meyerhof carbon flux pathway comprising a  
CC gene coding a pyrophosphate dependent phosphofructokinase enzyme or a 16s  
CC RNA. The bacterial strain is useful for the production of single cell  
CC protein and for the bioconversion of a nitrogen-containing compound,  
CC e.g. ammonia, nitrate, nitrite or nitrogen. It is also useful for the  
CC production of a feed product comprising a protein, carbohydrates and a  
CC pigment and for reducing oxygen demand, for removing nitrates and  
CC nitrates in methane-containing environments such as landfills, waste  
CC water treatment systems or anywhere that methane, oxygen and nitrates are  
CC present. The bacterial strain of the invention can be used as a  
CC dentrifying agent for the conversion of nitrate or nitrite to nitrous  
CC oxide with methane or methanol as a carbon source. It is also used in the  
CC production of biomass including proteins, carbohydrates and a wide  
CC variety of pigments (particularly for isoprenoid pigments for the  
CC purpose of generating animal feeds), in production of terpeneoid and  
CC carotenoid compounds, useful as pigments and as monomers in polymeric  
CC materials and in production of exopolysaccharides at high levels.  
CC Sequences ABG61551-ABG61590 represent high growth methanotrophic  
CC bacterial strain proteins of the invention.

XX Sequence 231 AA:

Query Match 100.0%; Score 1199; DB 23; Length 211;  
Best Local Similarity 100.0%; Pred. No. 5,7e-118;  
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPITQCAVAVPAAGVGRKQMDRKOYLPAGKVIHTLTRLLIESAPFQKVAVALSVE 60  
DB 1 MNPITQCAVAVPAAGVGRKQMDRKOYLPAGKVIHTLTRLLIESAPFQKVAVALSVE 60  
QY 61 DYPWPELSIAKHPDITITAPGKERADSVLSALKALEDIASENDWLVVDARPCITGSDI 120  
DB 61 DYPWPELSIAKHPDITITAPGKERADSVLSALKALEDIASENDWLVVDARPCITGSDI 120  
QY 121 HQIDITLKNDPVGGILALSSHDTLKHVGDITITATIDRKHWRALTPQMFYGMRLDALQ 180  
DB 121 HQIDITLKNDPVGGILALSSHDTLKHVGDITITATIDRKHWRALTPQMFYGMRLDALQ 180  
QY 181 RTEGNPAVTVDEASALELIGHKPKIVGGRPDNIKITRPBDLALAQFYEQQA 231  
DB 181 RTEGNPAVTVDEASALELIGHKPKIVGGRPDNIKITRPBDLALAQFYEQQA 231

RESULT 2

AAE22303

ID AAE22303 standard; Protein; 231 AA.

XX AAE22303;

XX 25-UTL-2002 (first entry)

XX Methylomonas 16a sp. 2C-methyl-d-erythritol cytidyltransferase enzyme.

XX Carotenoid; isopentenyl pyrophosphate; anthraxanthin; astaxanthin; diet;

XX anti-oxidant; steroid; flavour; fragrance; electro-optic application;

XX aquaculture; enzyme; 2C-methyl-d-erythritol cytidyltransferase; ispd.

XX Methylomonas 16a sp.

XX WO200218617-A2.

XX 07-MAR-2002.

PF 04-SEP-2001; 2001MO-US27420.

PR 01-SEP-2000; 2000US-229858P.

PR 01-SEP-2000; 2000US-229907P.

XX (DUPC) DU POINT DE MEMOIRS & CO E I.

PI Brzostowicz PC, Cheng Q, Dicosimo DJ, Koffas M, Miller ES;

PI Odum JM, Picotaggio SK, Rouviere PR;

DR WPI; 2002-351711/38.

DR N-PSDB; MAD35501.

PT Producing carotenoid compounds e.g. anthraxanthin and astaxanthin, by

PT using microorganisms having a nucleic acid molecule encoding enzymes in

PT the carotenoid biosynthetic pathway and which metabolize single carbon

PT substrates -

PS Claim 43; Page 115; 156pp; English.

XX The invention relates to a method for producing carotenoid compounds.  
CC The method comprises a transformed metabolising host cell, comprising  
CC suitable levels of isopentenyl pyrophosphate and a nucleic acid molecule  
CC encoding an enzyme in the carotenoid biosynthetic pathway, under the  
CC control of regulatory sequences, and contacting the host cell with carbon  
CC substrate to produce a carotenoid compound. The method is useful for  
CC producing carotenoid compounds such as anthraxanthin and astaxanthin, by  
CC using microorganisms having a nucleic acid molecule encoding enzymes in  
CC the carotenoid biosynthetic pathway and which metabolize single carbon  
CC substrates. The carotenoids have potent anti-oxidant properties useful in  
CC diet, and aquaculture elements. The carotenoids are also useful as  
CC intermediates in the synthesis of steroids, flavours and fragrances and  
CC compounds for potential electro-optic applications. The present sequence  
CC is Methylomonas 16a sp. 2C-methyl-d-erythritol cytidyltransferase (ispd)  
CC enzyme used in the invention.

XX Sequence 231 AA:

QY Query Match 100.0%; Score 1199; DB 23; Length 211;  
Best Local Similarity 100.0%; Pred. No. 5,7e-118;  
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPITQCAVAVPAAGVGRKQMDRKOYLPAGKVIHTLTRLLIESAPFQKVAVALSVE 60  
DB 1 MNPITQCAVAVPAAGVGRKQMDRKOYLPAGKVIHTLTRLLIESAPFQKVAVALSVE 60  
QY 61 DYPWPELSIAKHPDITITAPGKERADSVLSALKALEDIASENDWLVVDARPCITGSDI 120  
DB 61 DYPWPELSIAKHPDITITAPGKERADSVLSALKALEDIASENDWLVVDARPCITGSDI 120  
QY 121 HQIDITLKNDPVGGILALSSHDTLKHVGDITITATIDRKHWRALTPQMFYGMRLDALQ 180  
DB 121 HQIDITLKNDPVGGILALSSHDTLKHVGDITITATIDRKHWRALTPQMFYGMRLDALQ 180  
QY 181 RTEGNPAVTVDEASALELIGHKPKIVGGRPDNIKITRPBDLALAQFYEQQA 231  
DB 181 RTEGNPAVTVDEASALELIGHKPKIVGGRPDNIKITRPBDLALAQFYEQQA 231

RESULT 3

AAU80327

ID AAU80327 standard; Protein; 231 AA.

XX AAU80327;

XX 15-UTL-2002 (first entry)

XX Methylomonas 16a ORF3 ysbp/ispd protein sequence.

XX Isoprenoid biosynthetic enzyme; isoprenoid compound; feed additive;

XX carotenoid; pigment; flavour; fragrance; open reading frame 3; ORF3;

XX ysbp; 2C-methyl-D-erythritol cytidyltransferase enzyme; ispd.

OS Methylomonas sp.  
XX WO200220733-A2.  
XX  
XX 14-MAR-2002.  
XX  
XX 29-AUG-2001; 2001WO-US26852.  
XX  
XX 01-SEP-2000; 2000US-2299072.  
XX  
XX (DUPO) DU PONT DE NEMOURS & CO E. I.  
XX  
XX Cheng Q, Koffas M, Norton KC, Odom JM, Picataggio SK, Rouviere PE,  
XX Schenzle A, Tomb J,  
XX WPI: 2002-383051/41.  
XX N-PSDB; AAKS0083.  
XX  
XX Novel nucleic acid molecule encoding a isoprenoid biosynthetic enzyme,  
XX isolated from Methylomonas 16a, useful for the production of isoprenoid  
XX compounds -  
XX  
XX Claim 4; Page 70-71; 84pp; English.  
XX  
XX The present invention relates to a new nucleic acid molecule encoding  
XX an isoprenoid biosynthetic enzyme isolated from Methylomonas 16a.  
XX The invention is useful for obtaining a nucleic acid molecule  
XX encoding an isoprenoid compound biosynthetic enzyme, and for the  
XX microbial production of isoprenoid compounds. The molecules of the  
XX invention are also useful for regulating isoprenoid biosynthesis in an  
XX organism and for producing recombinant organisms for producing various  
XX isoprenoid compounds. The nucleic acid is also useful for feed additive,  
XX for the production of keratoids and their derivatives, isoprenoid  
XX intermediates, and as pure products useful as pigments, flavours and  
XX fragrances. The present amino acid sequence represents the Methylomonas  
XX 16a open reading frame 3 (ORF3) ygbP/ispD (2C-methyl-D-erythritol  
XX cytidyltransferase enzyme) protein of the invention, as described above.  
XX  
XX Sequence 231 AA;  
SQ  
Query Match 100.0%; Score 1199; DB 23; Length 231;  
Best Local Similarity 100.0%; Pred. No. 5,7e-118;  
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNPITQCAVAVPAAGVGKMQADRPKQYLPLAGKTVIEHTYTRLLSDFAPQKVAVALSVB 60  
DB 1 MNPITQCAVAVPAAGVGKMQADRPKQYLPLAGKTVIEHTYTRLLSDFAPQKVAVALSVB 60  
QY 61 DRYWDELSTIAKHPDITTAAGGERADSVLSALKLEDIASENDVAVHDAARPCLTGSDI 120  
DB 61 DRYWDELSTIAKHPDITTAAGGERADSVLSALKLEDIASENDVAVHDAARPCLTGSDI 120  
QY 121 HQIDTILKNDPVGGIILALSSHDITLKHVVDGTTITATIDRKHWRAALTPQMKYGMRLDALQ 180  
DB 121 HQIDTILKNDPVGGIILALSSHDITLKHVVDGTTITATIDRKHWRAALTPQMKYGMRLDALQ 180  
QY 181 RTEGNPAVTDASALELGHKPKIVEGRPDNIXITREPDALAPFYMEQQA 231  
DB 181 RTEGNPAVTDASALELGHKPKIVEGRPDNIXITREPDALAPFYMEQQA 231  
RESULT 4  
AAU36338 standard; Protein; 234 AA.  
XX  
XX AAU36338;  
XX  
XX 14-FEB-2002 (first entry)  
XX  
XX Pseudomonas aeruginosa cellular proliferation protein #328.  
XX  
XX Antisense; prokaryotic cellular proliferation protein;  
XX antibiotic; antibacterial; drug design.  
XX

XX  
XX OS pseudomonas aeruginosa.  
XX  
XX PN 0200170955-A2.  
XX  
XX PD 17-SEP-2001.  
XX  
XX 1-MAR-2001; 2001WO-US09180.  
XX  
XX 1-MAR-2000; 2000US-191078P.  
XX  
XX 3-MAY-2000; 2000US-206848P.  
XX  
XX 5-MAY-2000; 2000US-207727P.  
XX  
XX 3-OCT-2000; 2000US-242578P.  
XX  
XX 7-NOV-2000; 2000US-253625P.  
XX  
XX 2-DEC-2000; 2000US-257931P.  
XX  
XX 5-FEB-2001; 2001US-269308P.  
XX  
XX ELIT-1) ELITRA PHARM INC.  
XX  
XX Maselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,  
XX Samamoto RT, Xu HH;  
XX PI: 2001-611495/70.  
XX  
XX 1-PSDB; AAKS4197.  
XX  
XX 1-PI; 2001-611495/70.  
XX  
XX New polynucleotides for the identification and development of  
XX antibiotics, comprise sequences of antisense nucleic acids -  
XX  
XX Example 3; Seq ID No 11931; 511pp; English.  
XX  
XX The invention relates to antisense inhibitors of genes essential to  
XX prokaryotic cellular proliferation, their use in identifying the  
XX genes, their use in the discovery of novel antibiotics, the essential  
XX genes themselves and the encoded proteins. The prokaryotes used are  
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
XX invention is also useful for the identification of potential new targets  
XX for antibiotic development. The antisense nucleic acids can also be used  
XX to identify proteins used in proliferation, to express these proteins,  
XX and to obtain antibodies capable of binding to the expressed proteins.  
XX The proteins can be used to screen compounds in rational drug discovery  
XX programmes. The antisense nucleic acid sequence is also useful to screen  
XX for homologous nucleic acids which are required for cell proliferation in  
XX a wide variety of organisms. The present sequence represents an  
XX essential prokaryotic cellular proliferation protein.  
XX Note: The sequence data for this patent did not form part  
XX of the printed specification, but was obtained in electronic  
XX format directly from WIPO at  
XX [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).  
XX  
XX Sequence 234 AA;  
SQ  
Query Match 46.3%; Score 555.5; DB 22; Length 234;  
Best Local Similarity 54.1%; Pred. No. 4.9e-50;  
Matches 119; Conservative 26; Mismatches 72; Indels 3; Gaps 3;  
QY 8 KAVVPAAGVGKMQADRPKQYLPLAGKTVIEHTYTRLLSDFAPQKVAVALSVBPPPEL 67  
DB 10 KAVVPAAGVGKMQADRPKQYLPLAGKTVIEHTYTRLLSDFAPQKVAVALSVBPPPEL 69  
QY 68 SLAKHPDITTAAGGERADSVLSALKLEDIASENDVAVHDAARPCLTGSDITLQIDT 126  
DB 70 DCAARHVRQRAAGGKRRGVSILNGLRLERGAQADQVAVHDAARPCLTGSDITLQIDT 129  
QY 127 LKNDPVGGIILALSSHDITLKHVVDGTTITATIDRKHWRAALTPQMKYGMRLDALQ 184  
DB 130 LAEDPVGGIILAVPADITLKRSDRGVSEITDRSVVLAATPQMKYGMRLDALQ 189  
QY 185 NPVATDEASALELGHKPKIVEGRPDNIXITREPDALAPFYMEQQA 224  
DB 190 GVAITDEASALELGHKPKIVEGRPDNIXITREPDALAPFYMEQQA 229

RESULT 5  
AAU35483  
ID AAU35483 standard; Protein; 225 AA.  
XX  
XX AAU35483;  
AC  
XX  
XX  
DT 14-FEB-2002 (first entry)  
DE Haemophilus influenzae cellular proliferation protein #124.  
XX  
XX  
XX Actisense; prokaryotic cellular proliferation protein;  
XX antibiotic; antibacterial; drug design.  
OS Haemophilus influenzae.  
XX  
XX WO200170955-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 21-MAR-2001; 2001WO-US09180.  
PF  
XX  
XX 21-MAR-2000; 2000US-191078P.  
PR 23-MAY-2000; 2000US-206848P.  
PR 26-MAY-2000; 2000US-207727P.  
PR 23-OCT-2000; 2000US-242578P.  
PR 27-NOV-2000; 2000US-253625P.  
PR 22-DEC-2000; 2000US-257931P.  
PR 16-FEB-2001; 2001US-269308P.  
XX  
XX (ELIT- ) ELITRA PHARM INC.  
PA  
XX  
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HE;  
PI  
XX WPI; 2001-611495/70.  
DR N-PSDB; AAS53342.  
XX  
XX  
PT New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids -  
XX  
XX  
PS Example 3; Seq ID No 11076; 511pp; English.  
XX  
XX  
XX The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the  
CC genes, their use in the discovery of novel antibiotics, the essential  
CC genes themselves and the encoded proteins. The prokaryotes used are  
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
CC invention is also useful for the identification of potential new targets  
CC for antibiotic development. The antisense nucleic acids can also be used  
CC to identify proteins used in proliferation, to express these proteins,  
CC and to obtain antibodies capable of binding to the expressed proteins.  
CC The proteins can be used to screen compounds in rational drug discovery  
CC programmes. The antisense nucleic acid sequence is also useful to screen  
CC for homologous nucleic acids which are required for cell proliferation in  
CC a wide variety of organisms. The present sequence represents an  
CC essential prokaryotic cellular proliferation protein.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 225 AA;  
Query Match 44.2%; Score 530.5; DB 22; Length 225;  
Best Local Similarity 52.5%; Pred. No. 2e-47; Mismatches 65; Indels 13; Gaps 6;  
Matches 117; Conservative 28; Mismatches 65; Indels 13; Gaps 6;  
QY 9 AAVPAAGVGGKQWQADRPQYLTPLAGTIVIEHTLTRELSDAFORVAVAISVEDEPYWELLS 68  
DB 7 AVLPAGVGGKQWQADRPQYLTPLAGTIVIEHTLTRELSDAFORVAVAISVEDEPYWELLS 66  
QY 69 IAKHDIITAPGGRKADSVLSALKALDINSNDWVYVHNAARPCULTGSDIHQIDTLK 128

DB 67 L--DEKIQVSGGTTTAAESVLANGANA---IASKNAPVYVHNAARPCULGSDIHQIDTLK 117  
QY 129 --NDEVGIIALISHTLKHVDG--DTTATIDRKHWPRALTTPQFKRGMLDALQR--TEG 184  
DB 118 AIEDKQGAIIAIPVDTITKADNOQCIKVEDRSQIWMQAMTPQFPVDIIDLALSTGICQ 177  
QY 185 NPAVTDEASALELGHKPKIVGPRPNKIKIRPEDLALAPRYM 227  
DB 178 GANITDEASATILAGFRPHLVAGRSNKLKVRPEDLALAEYTL 220  
RESULT 6  
AAU36129  
ID AAU36129 standard; Protein; 254 AA.  
XX  
XX AAU36129;  
AC  
XX  
XX 14-FEB-2002 (first entry)  
DT  
XX  
XX  
XX *Klebsiella pneumoniae* cellular proliferation protein #117.  
DE  
XX  
XX  
XX Actisense; prokaryotic cellular proliferation protein;  
XX antibiotic; antibacterial; drug design.  
OS *Klebsiella pneumoniae*.  
XX  
XX WO200170955-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 21-MAR-2001; 2001WO-US09180.  
PF  
XX  
XX 21-MAR-2000; 2000US-191078P.  
PR 23-MAY-2000; 2000US-206848P.  
PR 26-MAY-2000; 2000US-207727P.  
PR 23-OCT-2000; 2000US-242578P.  
PR 27-NOV-2000; 2000US-253625P.  
PR 22-DEC-2000; 2000US-257931P.  
PR 16-FEB-2001; 2001US-269308P.  
XX  
XX  
XX (ELIT- ) ELITRA PHARM INC.  
PA  
XX  
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HE;  
PI  
XX WPI; 2001-611495/70.  
DR N-PSDB; AAS53342.  
XX  
XX  
PT New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids -  
XX  
XX  
PS Example 3; Seq ID No 11722; 511pp; English.  
XX  
XX  
XX The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the  
CC genes, their use in the discovery of novel antibiotics, the essential  
CC genes themselves and the encoded proteins. The prokaryotes used are  
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
CC invention is also useful for the identification of potential new targets  
CC for antibiotic development. The antisense nucleic acids can also be used  
CC to identify proteins used in proliferation, to express these proteins,  
CC and to obtain antibodies capable of binding to the expressed proteins.  
CC The proteins can be used to screen compounds in rational drug discovery  
CC programmes. The antisense nucleic acid sequence is also useful to screen  
CC for homologous nucleic acids which are required for cell proliferation in  
CC a wide variety of organisms. The present sequence represents an  
CC essential prokaryotic cellular proliferation protein.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX

XX	Sequence	254 AA:
Query Match	44.2%; Score 529.5; DB 22; Length 254;	
Best Local Similarity	51.1%; Pred. No. 3e-47;	
Matches 114; Conservative	30; Mismatches 70; Indels 9; Gaps 5	
QY	9 AVVPAGVGRQADPRKCYLTPLAGKYITETFLRLLESADFOKVAVAISVEDPYMPEIS 68	
DB	33 AVPPAAGFGRNDTETPKQYISIGKNTILLENVAALLADAAVQVVAIVAPGRRRSQLP 92	
QY	69 IACHPDIITAPGGRKADSVLSALXLEDIASENDWVLYEDNAPRCUTGSDIHLCQIDTLX 128	
DB	93 LAQHPIQIVYVDGGAERADSVLNGDLQAL----PEAQWLYVDNAPRCQHODLSRLSLCS 148	
QY	129 NDVYGGILALSHDITAKYVD-GDT-ITATIDKRYKRALTYQMKYKGLMDALQR--TGS 184	
DB	149 TSRVYGGILAAPVDITMKRAEPKRTAIATIVDNDIMHATPQQLPRELAVDCTLRALNBS 208	
QY	185 NPATDASALTELGHKPIVEGRPDNKTIRPEPDALAAQYF 227	
DB	209 -ATITDASALLEYCGFHNPQVAGRADNKTIVTRPEPDALAAEYFL 250	

RESULT 7  
 AAM47964  
 ID AAM47964 standard; Protein; 236 AA.  
 XX  
 AC AAM47964;  
 XX  
 DT 05-MAR-2002 (first entry)  
 XX  
 DE Escherichia coli CDP-ME synthase.  
 XX  
 KM Escherichia coli; ygdP; CDP-ME synthase; protein coordinate data;  
 KM 4-diphosphocytidyl-2-C-methylerythritol synthase; terpenoid; infection;  
 KM non-mevalonate isoprenoid; biosynthesis pathway; antibacterial; tetanus;  
 KM antidiarrheic; antiinflammatory; tuberculosal; Streptococcus; anthrax  
 XX toxic shock syndrome; meningitis; gonorrhea; gastroenteritis.  
 XX  
 OS Escherichia coli.  
 XX  
 PN MO200183769-A2.  
 XX  
 PD 08-NOV-2001.  
 XX  
 PF 03-MAY-2001; 2001MO-US14371.  
 XX  
 PR 03-MAY-2000; 2000US-201589P.  
 XX  
 PR 12-DEC-2000; 2000US-255088P.  
 XX  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX  
 PI Noel JP, Bowman ME, Richard S;  
 DR WPI; 2002-089742/12.  
 DR N-PSDS; ABR05928.  
 XX  
 PT Composition, useful for treating bacterial infections and for  
 PT identifying modulator compounds, comprise crystalline  
 XX 4-diphosphocytidyl-2-C-methylerythritol synthase -  
 XX  
 PS Example 1; Page -; 176pp; English.  
 XX  
 XX The invention relates to a composition (I) comprising CDP-ME,  
 CC 4-diphosphocytidyl-2-C-methylerythritol synthase in crystalline form. The  
 CC invention also discloses screening for compounds (II) that inhibit the  
 CC non-mevalonate isoprenoid biosynthesis pathway. (II) has antibacterial,  
 CC antidiarrheic, antiinflammatory and tuberculosal activity. (II) is  
 CC useful for inhibiting in vitro or in vivo, the activity of one or more  
 CC enzymes in the non-mevalonate isoprenoid biosynthesis pathway, in a cell  
 CC or cell-free environment, and thus modulating the growth of a cell e.g.  
 CC bacterial cell. (II) is also useful for inhibiting bacterial terpenoid

CC synthesis and treating a subject suffering from a bacterial infection  
CC of infection by *Streptococcus* or *Escherichia coli*. (II) is also useful  
CC for treating disorders caused by bacterial infections, including  
CC diarrhoea, pneumonia, dysentery, anthrax, typhoid fever, toxic shock  
CC syndrome, mastitis, meningitis, gonorrhoea, lymphoid fever,  
CC streptococcal, brucellosis, cholera, bubonic plague, tetanus,  
CC diphtheria and Lyme disease. The present sequence is that of the E.  
CC coli CPP-ME synthase.  
CC Note: The present sequence is not given in the specification but was  
CC generated from the encoding gene sequence at Genbank Accession Number  
CC U113362858.  
XX  
XX  
SO Sequence 236 AA;

Que / Match Similarity 43.7%; Score 524.5; DB 23; Length 236;  
 Res Local Similarity 50.2%; Pred. No. 9.2e-47;  
 Matches 112; Conservative 31; Mismatches 71; Indels 9; Gaps 4

QY 9 AVPPAAGVGKGMQADRPKQYIPLAGKTVIEHTLTLELSSDAFQKVAVAISVEDPYWPELS 68  
 DB 10 AVPPAAGGRRMGTCECPKQYLSIGNQTLLESHVALLAHPRVKRVVAISGDSRFQALP 69  
 QY 69 IAKGHDTITAPQGERADSVYSALKALIEDISENDMYVHQAARPCLTGSDIHQIDTLK 128  
 DB 70 LANHPQITVVDGSDRRADSVIAGLKA---AGDAQWVYHDAARPCLTQODLARIILASE 125  
 QY 129 NDPVGGLIALSSHDLTAKVD--GDITATITIDRKVWRALTEOMFYGMILADALQR--TEG 184  
 DB 126 TSNRGILIAAPVDIMKGAERQKNAIAITVVRNGIMHMLTQFPFRELHLDCLTALNEG 185  
 QY 185 NPATVDENSALELLGKHKVIVEGRPNDKITRPFEDLALAQRYM 227  
 DB 186 -AATDEASALEYCGFHQVWEGRADNIKVTRPFEDLALAEYL 227

RESULT 3  
 AB01172  
 ID AB01172 standard; protein; 236 AA.  
 XX AB011372;  
 DT 12-FEB-2001 (first entry)  
 XX  
 DE 1. coli YGHP protein.  
 XX  
 KW 1 JEB; KDTB; YGGF; YGGF; YHBC; YGBP; YGBB; YCHB; YCHB; antibacterial;  
 XX 1 treatment; infection.  
 XX  
 OS 1 Escherichia coli.  
 XX  
 FN 1 J19916176-A1.  
 PD 12-OCT-2000.  
 XX  
 PF 10-APR-1999; 99DB-1016176.  
 XX  
 PR 10-APR-1999; 99DB-1016176.  
 XX  
 PA 99DB ) BAYER AG.  
 PI Huetz H, Ehler K, Freiberg C, Spaltmann F, Wieland B;  
 PI 1abischinski H;  
 XX 1abischinski H;  
 DR 10-APR-2000-619611/62.  
 DR 1-PSDB; AAC66042.  
 XX  
 PT 1 essential genes from bacteria, useful in screening for antimicrobial  
 PT 1 genes, and related proteins, transformants and antisense sequences -  
 PS 1 disclosure; Page 21-22; 28pp; German.  
 CC 1 this invention describes novel Escherichia coli genes (I) encoding  
 CC 1 proteins (II) designated YGGF, YHBC, YGSD, YGBB, YCHB, YGBB, YGBB and



CC KDTB, and genes (Ia) that encode orthologous gene products (Iia) in  
 CC other microorganisms and which have antibacterial activity. Recombinant  
 CC microorganisms in which expression of (I) or (Ia) can be regulated are  
 CC used to identify compounds that bind to the gene products, particularly  
 CC in affinity selection assays. (Ii) and (Iia) are used to identify, or  
 CC prepare, antibodies and other proteins that bind to the gene products.  
 CC Substances that bind to (Ii) or (Iia) are potentially useful as  
 CC antibacterials for treating a wide range of infections in humans and  
 CC animals. Sequences antisense to (I) and (Ia) can also be used as  
 CC antibacterials. The specified genes are widely distributed in bacteria  
 CC but have no close homologs in eukaryotic cells.

CC  
 XX  
 SQ Sequence 236 AA;

Query Match 43.4%; Score 520.5; DB 21; Length 236;  
 Best Local Similarity 49.8%; Pred. No. 2,4e-46;  
 Matches 111; Conservative 31; Mismatches 72; Indels 9; Gaps 4;

QY 9 AVPAAGVGRKQMDRPPQYVPLAGKVIHTLRLIESDAPKVAIVSEDPYMEIS 68  
 DB 10 AVPAAGGRMRQTECPQYISIGNQILHSVALLAHRVKRVVIAISPGSRFAQLP 69  
 QY 69 IAKPDIITAPGKERADSVLSAKALEDIASENDVYLVDAAAPCLTGSDIHQIDTLK 128  
 DB 70 LANHPQITVVDGGERADSVLAGLKA-----AGDAQWLVLDAAAPCLHQDILARLALSE 125  
 QY 129 NDPVGITLSSHDTLKVD--GDTTATIDRKHWRALTPQWKYGMALDQLQ--TEG 184  
 DB 126 TSTRGTILAPVRDTPRAEGRKNAIHTVDRNGLWMLALPQFPREILHDCLTRALNBS 185  
 QY 185 NPATVDEASALELHGKPKIVEGRPDNIKITRPEDLALAQFM 227  
 DB 186 -ATTIDEASALEYCGFHPQVVEGRADNIKVTRPEDLALAEFYL 227

RESULT 9  
 AAU34680  
 ID AAU34680 standard; Protein; 236 AA.

AC AAU34680;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE E. coli cellular proliferation protein #261.  
 XX  
 KM Antisense, prokaryotic cellular proliferation protein;  
 KM antibiotic; antibacterial; drug design.  
 OS Escherichia coli.  
 XX  
 PN WO200170955-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PP 21-MAR-2001; 2001WO-US09180.  
 XX  
 PR 21-MAR-2000; 2000US-191078P.  
 PR 23-MAY-2000; 2000US-206848P.  
 PR 26-MAY-2000; 2000US-207227P.  
 PR 23-OCT-2000; 2000US-243578P.  
 PR 27-NOV-2000; 2000US-253625P.  
 PR 22-DEC-2000; 2000US-257931P.  
 PR 16-FEB-2001; 2001US-269308P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX  
 DR WPI; 2001-611495/70.  
 DR N-PSDB; AAS52539.  
 XX  
 PT New polynucleotides for the identification and development of

PR antibiotics, comprise sequences of antisense nucleic acids -  
 XX  
 XX Example 3; Seq ID No 10273; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC Escherichia coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
 CC pneumoniae, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pcr\_sequences.

SQ Sequence 236 AA;

QY Query Match 43.4%; Score 520.5; DB 22; Length 236;  
 Best Local Similarity 49.8%; Pred. No. 2,4e-46;  
 Matches 111; Conservative 31; Mismatches 72; Indels 9; Gaps 4;

QY 9 AVPAAGVGRKQMDRPPQYVPLAGKVIHTLRLIESDAPKVAIVSEDPYMEIS 68  
 DB 10 AVPAAGGRMRQTECPQYISIGNQILHSVALLAHRVKRVVIAISPGSRFAQLP 69  
 QY 69 IAKPDIITAPGKERADSVLSAKALEDIASENDVYLVDAAAPCLTGSDIHQIDTLK 128  
 DB 70 LANHPQITVVDGGERADSVLAGLKA-----AGDAQWLVLDAAAPCLHQDILARLALSE 125  
 QY 129 NDPVGITLSSHDTLKVD--GDTTATIDRKHWRALTPQWKYGMALDQLQ--TEG 184  
 DB 126 TSTRGTILAPVRDTPRAEGRKNAIHTVDRNGLWMLALPQFPREILHDCLTRALNBS 185  
 QY 185 NPATVDEASALELHGKPKIVEGRPDNIKITRPEDLALAQFM 227  
 DB 186 -ATTIDEASALEYCGFHPQVVEGRADNIKVTRPEDLALAEFYL 227

RESULT 10  
 AAB3 578  
 ID AAB31678 standard; Protein; 236 AA.

AC AAB31678;  
 XX  
 DT 07-MAR-2003 (first entry)  
 XX  
 DB Escherichia coli ybbp protein.  
 XX  
 DR 3-YaccM protein; 5-YaccJ protein; pharmaceutical formulation;  
 KM bacterial infection; antibacterial.  
 OS Escherichia coli.  
 XX  
 PN WO200281652-A2.  
 XX  
 PD 17-OCT-2002.  
 XX  
 PP 21-FEB-2002; 2002WO-US05086.  
 XX  
 PR 23-FEB-2001; 2001US-0792251.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Fritz C, Youngman P, Guzman L;





Db 559 -ATITDEASALEYCG 572

RESULT 14

AAV74692

ID AAV74692 standard; Protein; 229 AA.

XX AAV74692;

XX AAV74692;

DE 21-MAR-2000 (first entry)

DB Neisseria gonorrhoeae ORF 233 protein sequence SEQ ID NO:858.

XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

KM antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;

KM antibacterial; gene therapy.

XX Neisseria gonorrhoeae.

OS

PN W09957280-A2.

XX

PD 11-NOV-1999.

XX

PF 30-APR-1999; 99WC-US09346.

XX

PR 01-MAY-1998; 98US-0083758.

PR 31-JUL-1998; 98US-0094869.

PR 02-SEP-1998; 98US-0098894.

PR 02-SEP-1998; 98US-0099062.

PR 09-OCT-1998; 98US-0103749.

PR 09-OCT-1998; 98US-0103794.

PR 09-OCT-1998; 98US-0103796.

XX 25-FEB-1999; 98US-0121528.

PA (CHIR ) CHIRON CORP.

PA (GENO-) INST GENOMIC RES.

PI Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M,

PI Petersen J, Piza M, Rappoli R, Ratti G, Scalato E, Scarselli M,

PI Tetteien H, Venter JC;

XX WPI: 2000-062150/05.

DR N-PSDB; AAZ53454.

XX

XX Novel Neisserial polypeptides predicted to be useful antigens for

XX vaccines and diagnostics -

XX

XX Claim 2; Page 532-533; 1453bp; English.

XX

XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAV74253 to AAV75941

XX represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides

XX and polypeptides. AAZ54557 to AAZ54576 and AAZ54616 to AAZ55473 represent

XX PCR primers used in the exemplification of the present invention. The

XX polypeptides, the polynucleotides, antibodies and compositions of

XX the invention can be used as vaccines, as diagnostic reagents, and as

XX immunogenic compositions. The polypeptides can be used in the

XX manufacture of medicaments for treating or preventing infection due to

XX Neisseria bacteria (e.g. meningitis and septicaemia), to detect the

XX presence of Neisseria bacteria or to raise antibodies. They may also

XX be used to screen for agonists or antagonists, which may themselves

XX have use as antibacterial agents. The polynucleotides of the invention

XX may also be used in gene therapy protocols.

XX

XX Sequence 229 AA;

XX

XX Query Match 33.2%; Score 397.5; DB 21; Length 229;

XX Best Local Similarity 39.0%; Pred No. 2.2e-33;

XX Matches 89; Conservative 47; Mismatches 75; Indels 17; Gaps 5

XX

XX 9 AAVPAAGVGKQADRPKQYLFPLAGKTYIERTLRLLESDFQCAVAIVASVEDPY----- 63

XX 7 ALTPAAGIGRFCAADPKQYVEIGSKTYLIEHTLGIFFERHAEVDFVTVVVSPEDETPADKQ 66

QY	6	--WPELSIAKHPOITAPGKERADSVSL-KALE-DIASENDMVLVHDAARPCLTGSD	119
DB	67	TAFQVRVAKN-----GGTRAFETVNGVAKLELTGALFETDNLIVHDAARCLPSEA	119
QY	120	THLQIDTLKNDPVGGILALSSHDLLKVDGDTTATIDRKHVMFALTPQWFKYGMEDAL	179
DB	120	LALFISAGAAEGGILAFVADTLKRAESQISATVDRSGMLWAOQTFQFQAGLHRL	179
QY	180	QREGNPAVTDASALELGHKPKIVEAPDNKIKTRPEDALAQFYA 227	
DB	180	AAENNGITDERSAVKLGVRPLLIQDANLKLTPQDAIYIRLL 226	
RESULT	15		
ABP79	33		
ID	ABP79613	standard; Protein; 229 AA.	
XX	ABP79613:		
XX	AC		
DT	7-MAR-2003	(first entry)	
XX			
DE	gonorrhoeae amino acid sequence SEQ ID 5756.		
XX			
KX	antibacterial; infection; vaccine; gene therapy.		
OS	Neisseria gonorrhoeae.		
FN	12020279243-A2.		
PD	1-OCT-2002.		
XX			
BF	7-FEB-2002; 2002WO-1B02069.		
XX			
FR	13-FEB-2001; 2001GB-0003424.		
PA	CHIR-1) CHIRON SPA.		
PI	Montana MR, Pizza M, Masignani V, Monaci B;		
DR	131; 2003-058415/05.		
XX	PSDS; ABZ40583.		
PT	A protein from Neisseria gonorrhoeae, useful for the manufacture of a		
PS	medicament for treating or preventing N. gonorrhoeae infection		
XX			
XX	disclosure; Page 608; 815pp; English.		
XX			
CC	The present invention relates to proteins from Neisseria gonorrhoeae.		
CC	so disclosed are the nucleic acid molecules encoding the proteins and		
CC	antibodies that specifically bind to the proteins. The composition		
CC	comprising the protein, nucleic acid or antibody is useful for the		
CC	manufacture of a medicament for treating or preventing N. gonorrhoeae		
CC	infection, this may be in the form of a vaccine or gene therapy.		
CC	sequences given in records ABP76736-ABP81046 represent nucleic acid		
CC	sequences of the invention.		
CC			
CC	sequence 229 AA;		
Que / Match	33.2%;	Score 397.5; DB 24; Length 229;	
Best Local Similarity	39.0%;	Pred. No. 2,2e-33;	
Matches	89; Conservative	47; Mismatches 75; Indels 17; Gaps 5	
QY	9	AYVPAAGVGRNQADRPKOYELPLAGKIVYIHTLRLBSDAFQVAVAIYVEDPY----- 63	
DB	7	ALIPAPAGIGRFGAPKPKOYVEIGSKYLLHRYLGIFFERHEAVDVLTVVAVSPEDTFADKQ 66	
QY	64	--WPELSIAKHEDITAPGKERADSVSL-KALE-DIASENDMVLVHDAARPCLTGSD 119	
DB	67	TAFQVRVAKN-----GGTRAFETVNGVAKLELTGALFETDNLIVHDAARCLPSEA 119	
QY	120	THLQIDTLKNDPVGGILALSSHDLLKVDGDTTATIDRKHVMFALTPQWFKYGMEDAL 179	

Thu Jan 29 17:41:17 2004

us-09-941-947a-10 pag

Page 10

Db 120 IARLEOAGNAEAGGILAVPVADTLKRAESGOISATVDRSGIMQAOPTPOLFOAGLIRAL 179

Oy 180 QRTGPNPAVDEASALELIGHKPKIVEGRPNIKITRPEDLALAQFYW 227

Db 180 -AAENUGGITDEASAVENKGVFRPLLIOGDARNEKLTQPDAYIVRLIL 226

Search completed: January 29, 2004, 15:49:40  
Job time : 22.1153 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 15:45:34 : Search time 7.2499 Seconds  
(without alignments)  
1348.130 Million cell updates/sec

Title: US-09-941-947a-10  
Perfect score: 1199

Sequence: 1 MNPTICQMAVPAAGVCKM.....IKITRPEDLALAQFYEQQA 231

Scoring table: ELOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA+  
1: /cgn2\_6/ptodata/1/aa/5A COMB pep: +  
2: /cgn2\_6/ptodata/1/aa/5B COMB pep: +  
3: /cgn2\_6/ptodata/1/aa/6A COMB pep: +  
4: /cgn2\_6/ptodata/1/aa/6B COMB pep: +  
5: /cgn2\_6/ptodata/1/aa/PTCUS COMB pep: +  
6: /cgn2\_6/ptodata/1/aa/backfiletest.pep: +

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	555.5	46.3	270	US-09-252-991A-21226	Sequence 21226, A
2	414.5	34.6	242	US-09-328-352-7400	Sequence 7400, App
3	240.5	20.1	218	US-09-198-452A-621	Sequence 621, App
4	237.5	19.8	238	US-09-107-532A-5165	Sequence 5165, App
5	218.5	18.2	271	US-09-134-001C-5299	Sequence 5299, App
6	208	17.3	235	US-09-285-576-2	Sequence 2, App1
7	104	8.7	461	US-09-328-352-5107	Sequence 5107, App
8	93	7.8	357	US-09-198-452A-487	Sequence 487, App
9	92	7.5	393	US-09-107-532A-6010	Sequence 6010, App
10	89.5	7.5	285	US-08-311-731A-34	Sequence 34, App1
11	89.5	7.5	285	US-08-311-731A-36	Sequence 36, App1
12	89	7.4	154	US-08-936-165A-528	Sequence 528, App
13	89	7.4	531	US-09-134-001C-5574	Sequence 5574, App
14	88.5	7.4	459	US-08-971-782-2	Sequence 2, App1
15	88.5	7.4	459	US-09-109-026-2	Sequence 2, App1
16	87.5	7.2	388	US-09-252-991A-17435	Sequence 17435, A
17	86.5	7.2	1788	US-08-962-284-2	Sequence 2, App1
18	83.5	7.0	398	US-09-303-064-54	Sequence 54, App1
19	83.5	7.0	398	US-09-086-503-54	Sequence 54, App1
20	83.5	7.0	405	US-09-252-991A-24036	Sequence 24036, A
21	83	6.9	344	US-08-867-611-37	Sequence 37, App1
22	83	6.9	344	US-08-867-611-38	Sequence 38, App1
23	83	6.9	375	US-08-867-611-50	Sequence 50, App1
24	83	6.9	375	US-08-500-222-6	Sequence 6, App1
25	83	6.9	375	US-08-500-125-6	Sequence 6, App1
26	83	6.9	376	US-07-779-704B-6	Sequence 6, App1
27	83	6.9	387	US-08-967-611-12	Sequence 12, App1

2	83	6.9	387	5	PCT-US92-06965A-17	Sequence 17, App1
2	83	6.9	393	3	US-08-867-611-14	Sequence 14, App1
3	83	6.9	393	5	PCT-US92-06965A-19	Sequence 19, App1
3	83	6.9	396	3	US-08-867-611-2	Sequence 2, App1
3	83	6.9	396	5	PCT-US92-06965A-7	Sequence 7, App1
3	83	6.9	417	3	US-08-867-611-20	Sequence 20, App1
3	83	6.9	417	5	PCT-US92-06965A-25	Sequence 25, App1
3	83	6.9	425	5	US-08-867-611-22	Sequence 22, App1
3	83	6.9	425	5	PCT-US92-06965A-27	Sequence 27, App1
3	83	6.9	467	3	US-08-867-611-24	Sequence 24, App1
3	83	6.9	467	5	PCT-US92-06965A-29	Sequence 29, App1
3	83	6.9	474	3	US-08-867-611-26	Sequence 26, App1
4	83	6.9	474	5	PCT-US92-06965A-31	Sequence 31, App1
4	83	6.9	491	2	US-08-912-129A-56	Sequence 56, App1
4	83	6.9	496	3	US-08-867-611-10	Sequence 10, App1
4	83	6.9	496	5	PCT-US92-06965A-15	Sequence 15, App1
4	83	6.9	498	1	US-08-500-222-2	Sequence 2, App1
4	83	6.9	498	1	US-08-500-125-2	Sequence 2, App1

#### ALIGNMENTS

RESULT 1  
US-09-252-991A-21226  
Sequence 21226, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUSINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
PI % REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252, 991A  
CURRENT FILING DATE: 1999-02-18  
PR OR APPLICATION NUMBER: US 60/074, 788  
PR OR FILING DATE: 1998-02-18  
PR OR APPLICATION NUMBER: US 60/094, 190  
PR OR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 21226  
LENGTH: 270  
T SE: PRT  
I SE: PRT  
C SE: PRT  
US-09-252-991A-21226  
Pseudomonas aeruginosa

Query Match 46.3%; Score 555.5; DB 4; Length 270;  
Best Local Similarity 54.1%; Pred. No. 2.5e-57;  
Matches 119; Conservative 26; Mismatches 72; Indels 3; Gaps 3;

QY 8 MAVVPAAGVCKMQRDRPROYLPLAGKTVIEHTLRLLESDFAPKAVAVISVDPYMPDL 67  
DB KVVTPAAGVCKMQRDRPROYLPLAGKTVIEHTLRLLESDFAPKAVAVISVDPYMPDL 105  
QY 68 STAKPDITAPGKERADSVLSA-LKALEDIASEMDVAVDAARPTLTSDFHLDIT 126  
DB DCAASHVGRAGAGRAAGVSVLNGHLLRLEFGAQDDVAVDAARPTLTSDFHLDIT 165  
QY 127 LKNDPVGGLASSHDLTKHNGD-TITATIDRKHVAVLTPQWVKYGMALD-LQRTBG 184  
DB LKNDPVGGLAVPADTLKSDRDRVSEITDRSVVAVLTPQWVKYGMALD-LQRTBG 225  
QY 185 NPATVDEASALTLGKRPVIEGRDNTKITRPEDLALAQ 224  
DB GVALTDEASAMEMAGVAPLVYGRADNTKITRPEDLALAQ 265  
RESULT 2  
US-09-328-352-7400  
Sequence 7400, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: GTC99-03BA  
 ; CURRENT APPLICATION NUMBER: US/09/328,352  
 ; CURRENT FILING DATE: 1999-06-04  
 ; NUMBER OF SEQ ID NOS: 8252  
 ; SEQ ID NO 7400  
 ; LENGTH: 242  
 ; TYPE: PR1  
 ; ORGANISM: Acinetobacter baumannii  
 US-09-328-352-7400

Query Match 34.6%; Score 414.5; DB 4; Length 242;  
 Best Local Similarity 43.1%; Pred. No. 1e-40;  
 Matches 97; Conservative 34; Mismatches 91; Indels 3; Gaps 3;

QY 8 WAVPAGVGRMADRPKQYPLAGKTVIEHTLRLLESDAFQKVAIVSDEPYWPEL 67  
 DB 18 WAVPAGSGSRFSEKTELKQYIQDAIVLEHVRLSOLPLGTYV-LALGKODTPASTL 76  
 QY 68 SIAGHPDIIITAPGKERADSVLSALKALEDIASENDVWLVHDAARPCITGSDIHLQIDTL 127  
 DB 77 SFQDHRKAFHFCNGGVERHVSALNVLNLSQIADDDVVLVHADARPCVTEGCLNTLVKNA 136  
 QY 128 KNDPVGGILASSHDTLKHVDG-D-TTATIDRKAVRALTPQFKYGMRLDALT-EGN 185  
 DB 137 IETVQSALIAIPVBDTLKQVNOEQOIKTVSRRLMQAOTPOLAKTILKKALETALKN 196  
 QY 186 PAVTDEASALELHGKPKIVEGRPNIKITPEPDLAAQFMEQO 230  
 DB 197 LTIIDFASALSIGSVQVWGRDNIKITPPDLLEAARLLSQO 241

RESULT 3  
 US-09-198-452A-621  
 ; Sequence 621, Application US/09198452A  
 ; Patent No. 6559294

; GENERAL INFORMATION:  
 ; APPLICANT: Grifflais, R.  
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
 ; TITLE OF INVENTION: and treatment of infection  
 ; FILE REFERENCE: 9710-003-999  
 ; CURRENT APPLICATION NUMBER: US/09/198,452A  
 ; CURRENT FILING DATE: 1998-11-24  
 ; NUMBER OF SEQ ID NOS: 6843  
 ; SEQ ID NO 621  
 ; LENGTH: 218  
 ; TYPE: PR1  
 ; ORGANISM: Chlamydia pneumoniae  
 US-09-198-452A-621

Query Match 20.1%; Score 240.5; DB 4; Length 218;  
 Best Local Similarity 32.9%; Pred. No. 3.6e-20;  
 Matches 74; Conservative 41; Mismatches 91; Indels 19; Gaps 9;

QY 3 PTICWAVPAGVGRMADRPKQYPLAGKTVIEHTLRLLESDAFQKVAIVSDEP 62  
 DB 7 PMIKSILILSGGQCTRFSGSKIPKQYPLNGTPIVLSL-KILSS-LPQIAEIVVCDP 63  
 QY 63 YWPELSIAKHDPDITAPGKERADSVLSALKALEDIASENDVWLVHDAARPCITGSDIHL 122  
 DB 64 SYQE-TFQETVPVSPAI-P-GERRQDSVFEGLOQV-----SYPMVVIHGGAPFIYPEIH- 115  
 QY 123 QIDTAKNDPVGGILASSHDTLKHVDGDTITATIDRKAVRALTPQFKYGMRLDALTQ 180  
 DB 116 --DLSTAEKIGATLAPIPYTIKQNP---VATIDRDNALIITFQCIKTELHAGLA 170  
 QY 181 -RTGEPNAVTDASALELHGKPKIVEGRPNIKITPEPDLAAQ 224  
 DB 171 IAKERQTLTVDDIEAARITIGKRSQIVFNKHPQIKISYPEDLITIQ 215

RESULT 4

US-0-107-532A-5165  
 ; Sequence 5165, Application US/09107532A  
 ; Patent No. 6583275

; GENERAL INFORMATION:  
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 7310  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: GENOME THERAPEUTICS CORPORATION  
 ; STREET: 100 Beaver Street  
 ; CITY: Waltham  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Arinello, Pamela Dereke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

FORMATION FOR SEQ ID NO: 5165:

SEQUENCE CHARACTERISTICS:

LENGTH: 218 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...238

SEQUENCE DESCRIPTION: SEQ ID NO: 5165:

US-0-107-532A-5165

QY Match 19.8%; Score 237.5; DB 4; Length 238;  
 Best Local Similarity 28.8%; Pred. No. 9.3e-20;  
 Matches 68; Conservative 52; Mismatches 95; Indels 21; Gaps 9;

QY 9 AVPAGVGRMADRPKQYPLAGKTVIEHTLRLLESDAFQKVAIVSDEPYWPEL 64  
 DB 9 AIIAGVGRMADRPKQYPLAGKTVIEHTLRLLESDAFQKVAIVSDEPYWPEL 68  
 QY 65 -PELSIAKHDPDITAPGKERADSVLSALKALEDIASENDVWLVHDAARPC---LTGSDI 120  
 DB 69 IEFNISVKKVIT--GSGKQSSINNGVQFLAKNYSDDDTITIHGIRPLVDELVSIV 126  
 QY 121 HLOIDTLKNDPVGGILASSHDT-LKHVDGDTITATIDRKAVRALTPQFKYGMRLDALTQ 178  
 DB 127 IYKCEFGN---GVTSIPVNRQFIKTE-ETTEQYVDNLTLRVSTPQAVYQYKLRRA 181  
 QY 179 LQR-TEGNPAVTDAA--SALLELHGKPKIVEGRPNIKITPEPDLAAQFYEQO 230  
 DB 182 YDRAVKEDIGMTSSVYNTVMVDGDTLTPAGSDKNIKITTTDBLELFAIVLKMK 237

RESULT 5  
 US-0-134-001C-5299

```

: Sequence 5299, Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: CURRENT FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 5299
: LENGTH: 271
: TYPE: RNT
: ORGANISM: Staphylococcus epidermidis
: US-09-134-001C-5299

```

Query Match	18.2%;	Score 218.5;	DB 4;	Length 271;
Best Local Similarity	29.6%;	Pred. No. 2e-17;		
Matches 72; Conservative	43;	Mismatches 83;	Indels 45;	Gaps 11;

```

QY      8  WAWPAPGAVCKEM-QADR-SKOYL-ELAGKVI-EHTLTLLIESPAFOKVAALSVEDPYWPE 66
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      36  YAGIAGLOGISRMGNVPLPKPQFISLOQGRKIIHHTYKFLMKDPBELLAL-----PQ 88
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      67  LSTAKHPDIT-----TAPGKERADSVLSALXAL-ED--IASENDVAVEDHARP 113
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      89  KWNMYMLDLNNYQLDDBKKIKYILOGGDDRNHSHIMNIIESIDCHKKADNEDIIVTHDVRP 148
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      114  CLTGSDDI-----HLOIDFLTKM--DVGWGLIALSSEHOTLKHVDEGDTTATIDRAGW 162
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      149  FLNRLIRREWEVYASQYGAVDITYNANVDIT-----ISSNDA-QFISGIFPI-----RSEMY 197
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      163  RALTPOWFKYGMRLDA-IQRTGKNPAV-TDEASALELLGHKPKVIEGRPNIKITRDEBL 220
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      198  QOGOTPCQFPEKIKELMDSYLSLTOSOKRILTDACKILVELEGKPVKLVGELFNKIKITTPYDL 257
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      221  ALA 223
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      258  KVA 260
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

```

RESULT 6
US-09-285-576-2
Sequence 2, Application US/09285576
Parent No. 6268177
GENERAL INFORMATION:
APPLICANT: Zalacain, Magdalena
APPLICANT: Burnham, Martin K. R.
APPLICANT: Brown, James R.
APPLICANT: Ingram, Karen I.
APPLICANT: Chalker, Alison F.
APPLICANT: Holmes, David T.
APPLICANT: Warren, Richard L.
APPLICANT: Mathie, Thomas B.
TITLE OF INVENTION: ndp
FILE REFERENCE: GM10167
CURRENT APPLICATION NUMBER: US/09/285,576
CURRENT FILING DATE: 1999-04-02
EARLIER APPLICATION NUMBER: 60/101,356
EARLIER FILING DATE: 1999-09-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 235
TYPE: PAT
ORGANISM: Streptococcus pneumoniae
US-09-285-576-2

```

Query Match 17.3%; Score 208; DB 3; Length 235;

Bes Local Similarity 27.6%; Pred. NO. 2.8e-16;  
 Mismatches 67; Conservative 48; Mismatches 96; Indels 32; Gaps 10.

```

Oy 8 WAWVPAAGVGKKNQ--ADRPKOTPLTLAGKVIETHITPLLESDAOCQVVAI-----57
Db 3 YAGIAGGCTGTNGISNLPKOTFLBJODRPILHTIEKFLVLEPSIEKIYVVGHGQWVSHAE 62
Oy 58 SVEDPYWPELSIAKHDDIITAPGKGERADSVLSALKALEDI--ASENDWVLVHDARPCL 116
Db 63 DEVDKTLIP---LYKERIITIT-KGADRNMTSICKIEALDAIDARPLTPBEDIVVTHDSVRPFI 116
Oy 116 TGSDDHLOIDTLKN-DPVGGIILASHHTL-KHNDGDTITATIKRXVWALTPOMER-- 177
Db 119 TLRMTDNTQLQANHDVAQTV--VEAVDTIVESTNGQFTIIPNRAHLYQOQTQOTFECK 176
Oy 116 TGSDDHLOIDTLKN-DPVGGIILASHHTL-KHNDGDTITATIKRXVWALTPOMER-- 177
Db 119 TLRMTDNTQLQANHDVAQTV--VEAVDTIVESTNGQFTIIPNRAHLYQOQTQOTFECK 176
Oy 172 -----YGMRLDALQRTGEGNPAVTDASALELLELGHKPKVBEGRPNIIKTRPEEDIALAQFY 226
Db 177 DPMDLYXGSLSD-----EKKELITDACKIFVIKKQKQVALAKKEYISNLKITTYTDTDKIAKSM 23

```

QY	227	MEQ	229
	:	:	:
Db	232	IEK	234

```

RESULT: 7
US-09 128-352-5107
; Seq: Ince 5107, Application US/09328352
; Patent No. 6562958
; GEN: AL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; PRI: REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5107
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09 128-352-5107

```

```

Que. / Match      8.7%; Score 104; DB 4; Length 461;
Res: Local Similarity 21.7%; Pred. NO. 0.0017;
Matches 64; Conservative 38; Mismatches 75; Indels 118; Gaps 16

```

```

QY 1KNPTICQAVVPAVGVGRMOADPKOYLPLAGVIVIEHTLTLESDAFQKVAIVSVE 60
Dh 8MSTTV---TILAAKGTIRMSQJLPKYOPLAGRPILGHV-----KTAQGLAAE 53
QY 61DPYWEPLSIAPHDITTPAGSKERADSVLSALKALEDIASEN-DWLVVHDAAAPCLTSD 119
Dh 54-----NITTVYG--HGGDHV-----KXTFAQENIQWV--EQAEOLGTGHA 89
QY 120IHKOLDTLKNDPVG---GILATSSHDTLK-----HYDGGT----- 151
Dh 90VQMTLPVLPKDCISLILYGDVPLVRCOTTLLEQOLIAENKGTIGIMTILHVDNPPYGGYIRYO 149
QY 152---TATIDRK-----HYWRAALTPQMEKXYGMLZDALQRTG 184
Dh 150DCKIQAIVESHKQATPAQRQIQEINTGIVYCSNAKLHEW---LPKL-----SNNNAQG 198
QY 185NRAVYD---BASATLTLGHKKPKI---VEGRBDNKITIRPBDLALACPYMEQQA 221
Dh 199EYTLDPVAMAVADGETASTIQPLAEVGVGVDRIQOLA---ALERFPQOQQA 248

```

RESULT: 8  
US-09-198-452A-487  
; Sequence 487, Application US/09198452A  
; Patent No. 6559294  
; GENERAL INFORMATION:  
; APPLICANT: Griffeais, R.



;; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
;; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
;; FILE REFERENCE: 9710-003-999  
;; CURRENT APPLICATION NUMBER: US/09/198,452A  
;; NUMBER OF SEQ ID NOS: 6849  
;; SEQ ID NO: 487  
;; LENGTH: 357  
;; TYPE: PRT  
;; ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-487

Query Match 7.8%; Score 93; DB 4; Length 357;  
Best Local Similarity 26.0%; Pred. No. 0.023;  
Matches 51; Conservative 30; Mismatches 79; Indels 36; Gaps 11;

QY 11 VPAAGVGRMO--ADPRQYL-----FLAGKVIHTLRLLSDAFQKVAVAISVEDPY 63  
DB 92 LPPEAVSQLELSDGYSYLVAVRCDSGETVQORL--LINDSGRSMTVVIS---B 145  
QY 64 WPELSIAGHPDI-----ITAPGKERADSVLSALKALEDIASENDVVLVHDARPC 114  
DB 146 LPE-----GHPDIRNLQSLASERIFVSRERADAVASCKVAFPDHILPWSSHIAVEE 201  
QY 115 LTGSDIHIQITLKNDPVGGILALSHDTLKHVDGDTITATIDPKHYWRALTPQMEKYG 174  
DB 202 IREKQQTMOGSLTSEDLGALLC-NTVSTETKTL-AFALDAVL-KQSVRRRNDLFAVE- 257  
QY 175 LRDALQRTGNPAVD 190  
DB 258 -RALE-----ASVTD 267

RESULT 9  
US-09-107-532A-6010  
; Sequence 6010, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:

;; APPLICANT: Lynn A Doucette-Stamm and David Bush  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
;; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
;; NUMBER OF SEQUENCES: 7310  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESS: GENOME THERAPEUTICS CORPORATION  
;; STREET: 100 Beaver Street  
;; CITY: Waltham  
;; STATE: Massachusetts  
;; COUNTRY: USA  
;; ZIP: 02354  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: CD-ROM ISO9660  
;; COMPUTER: PC  
;; OPERATING SYSTEM: <Unknown>  
;; SOFTWARE: ASCII  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/107,532A  
;; FILING DATE: 30-Jun-1998  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 60/085,596  
;; FILING DATE: 14 May 1998  
;; APPLICATION NUMBER: 60/051571  
;; FILING DATE: July 2, 1997  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Ariniello, Pamela Denise  
;; REGISTRATION NUMBER: 40,489  
;; REFERENCE/DOCKET NUMBER: GTC-012  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (781)893-5007  
;; TELEFAX: (781)893-8277  
;; INFORMATION FOR SEQ ID NO: 6010:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 393 amino acids

;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: YES  
;; ORIGINAL SOURCE:  
;; ORGANISM: Enterococcus faecium  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (B) LOCATION 1...393  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 6010:  
US-C-107-532A-6010

QY Match 7.7%; Score 92; DB 4; Length 393;  
Best Local Similarity 21.6%; Pred. No. 0.035;  
Matches 53; Conservative 30; Mismatches 72; Indels 90; Gaps 9;

QY 8 WAPVPAAGVGRMOADSPKQYLPLAGKVIHTLRLLSDAFQKVAVAISVEDPYWPEL 67  
DB 17 YAIILAAKGRTRMSKLYKYLHPVSGKPMTEHILNRVSEI----- 56  
QY 68 SIARHPDITAPG-GKERADSVL-----SALKALEDIASENDVVLVHD 109  
DB 57 ---KPEDEVIITVGHGAGQVKAQGERSKYALQAEQGTGAVALQASFLSKGGTTLVIS 113  
QY 110 AARPLTGSQDHLQIDTLKNDPVGGILALSHDTLKHVDGDTITATIDPKH 160  
DB 114 GDTPLLT-----TETLNN-----LREYHOGKNAATILTRQANPFGRIIRDH 158  
QY 161 V-----WRALTPQ-----MFYKGLRDALQRTGNPA-----VTDEASAL 195  
DB 159 IGIVEKIVEQDAPFEALVQGIINTGYCFQNEALFALSKVGTNNAGGYALTDILIEL 218  
QY 196 ELGGH 200  
DB 219 KEKGH 223

RESULT 10  
US-0-311-731A-34  
; Sequence 34, Application US/08311731A  
; Patent No. 6583266  
; GENERAL INFORMATION:

;; APPLICANT: SMITH, DOUGLAS  
;; APPLICANT: MAO, JEN-I  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
;; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR  
;; NUMBER OF SEQUENCES: 411  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
;; STREET: 600 ATLANTIC AVENUE  
;; CITY: BOSTON  
;; STATE: MASSACHUSETTS  
;; COUNTRY: USA  
;; ZIP: 02210  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/311,731A  
;; FILING DATE:  
;; CLASSIFICATION: 530  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: GATES, EDWARD R.  
;; REGISTRATION NUMBER: 31,616  
;; REFERENCE/DOCKET NUMBER: C0044/7125  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617/720-3500  
;; TELEFAX: 617/720-2441  
;; INFORMATION FOR SEQ ID NO: 34:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 393 amino acids

LENGTH: 285 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: Protein  
 HYPOTHEICAL: YES  
 ORIGINAL SOURCE:  
 ORGANISM: MYCOBACTERIUM LEPRAE  
 US-08-311-731A-34

Query Match 7.5%; Score 89.5; DB 4; Length 285;  
 Best Local Similarity 18.9%; Pred. No. 0.042;

Matches 55; Conservative 43; Mismatches 102; Indels 91; Gaps 9;

QY 3 PTTCMAVVP-----AAGYCKMQADRPQYIPLAGKYIETHTLRLLESDAFQ 51  
 DB 10 PTLMAAILADTSNRVWVCCGAGVGKTTTAAAMALQAAEYGRIVV-----VLTIDPAK 63  
 QY 52 KYAAVAISVED-----PYMPELSIAKH-----PDIITAPGKERADSVLS--- 90  
 DB 64 RLAAQLAGVNDLGNTPQRYPLAEPVPELHAMMLDMRRTFDEMNVQYSGPERAAQAILNSEF 123  
 QY 91 -----ALKALEDIASENDVIV---HDAAPCLTGSIDHILQIDTLKNDPV 132  
 DB 124 YQTVATSLAGTQEVNMEKLGQLLSQDRMDLVVDTPPSRNALDFLDAPKRLGNFMNSRL 183  
 QY 133 GGIILASSHDTLKHVDGDTTATIDRKHWNRALTPQMFYKGMALDALQRTGNDPAVTDEA 192  
 DB 184 GRLILTPG-----RGIGRLVTGAM---GIAMRALSTVLSGQMLADAA 222  
 QY 193 SALELLGHRKPKIVEGRPDN-----IKITRPEDIALAQFYMEQ 229  
 DB 223 TFVQSLDATFGSGFRKADRTYALTKRGTOFVWVSAAPDRLREASPFVDR 273

## RESULT 11

US-08-311-731A-36  
 Sequence 36, Application US/08311731A  
 Patent No. 6583266

GENERAL INFORMATION:

APPLICANT: SMITH, DOUGLAS

APPLICANT: MAO, JEN-I

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES

TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR

NUMBER OF SEQUENCES: 411

CORRESPONDENCE ADDRESS:

ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.

STREET: 600 ATLANTIC AVENUE

CITY: BOSTON

STATE: MASSACHUSETTS

COUNTRY: USA

ZIP: 02210

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/311,731A

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: GATES, EDWARD R.

REGISTRATION NUMBER: 31,616

REFERENCE/DOCKET NUMBER: C0044/7125

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/720-3500

TELEFAX: 617/720-2441

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 285 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: Protein  
 HYPOTHEICAL: YES  
 ORIGINAL SOURCE:  
 ORGANISM: MYCOBACTERIUM LEPRAE  
 US-08-311-731A-36

Query Match 7.5%; Score 89.5; DB 4; Length 285;  
 Best Local Similarity 18.9%; Pred. No. 0.042;

Matches 55; Conservative 43; Mismatches 102; Indels 91; Gaps 9;

QY 3 PTTCMAVVP-----AAGYCKMQADRPQYIPLAGKYIETHTLRLLESDAFQ 51  
 DB 10 PTLMAAILADTSNRVWVCCGAGVGKTTTAAAMALQAAEYGRIVV-----VLTIDPAK 63  
 QY 52 KYAAVAISVED-----PYMPELSIAKH-----PDIITAPGKERADSVLS--- 90  
 DB 64 RLAAQLAGVNDLGNTPQRYPLAEPVPELHAMMLDMRRTFDEMNVQYSGPERAAQAILNSEF 123  
 QY 91 -----ALKALEDIASENDVIV---HDAAPCLTGSIDHILQIDTLKNDPV 132  
 DB 124 YQTVATSLAGTQEVNMEKLGQLLSQDRMDLVVDTPPSRNALDFLDAPKRLGNFMNSRL 183  
 QY 133 GGIILASSHDTLKHVDGDTTATIDRKHWNRALTPQMFYKGMALDALQRTGNDPAVTDEA 192  
 DB 184 GRLILTPG-----RGIGRLVTGAM---GIAMRALSTVLSGQMLADAA 222  
 QY 193 SALELLGHRKPKIVEGRPDN-----IKITRPEDIALAQFYMEQ 229  
 DB 223 TFVQSLDATFGSGFRKADRTYALTKRGTOFVWVSAAPDRLREASPFVDR 273

## RESULT 12

US-08-936-165A-528

Sequence 528, Application US/08936165A  
 Patent No. 6348382

GENERAL INFORMATION:

APPLICANT: Black, Michael

APPLICANT: Burnham, Martin

APPLICANT: Hodgson, John

APPLICANT: Knowles, David

APPLICANT: Lometo, Michael

APPLICANT: Nicholas, Richard

APPLICANT: Pratt, Julie

APPLICANT: Reichberg, Richard

APPLICANT: Rosenberg, Martin

APPLICANT: Ward, Judith

TITLE OF INVENTION: No. 6348382e1 Prokaryotic Polynucleotides,

TITLE OF INVENTION: Polypeptides and Their Uses

NUMBER OF SEQUENCES: 534

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/936,165A

FILING DATE: 24-SEP-1997

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: GIMMI, EDWARD R.

REGISTRATION NUMBER: 38,891

REFERENCE/DOCKET NUMBER: P50549

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-4478  
 TELEFAX: 610-270-5090  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 528:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 154 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: Protein  
 US-08-936-155A-528

Query Match 7.4%; Score 89; DB 4; Length 154;  
 Best Local Similarity 23.9%; Pred. No. 0.018;  
 Matches 26; Conservative 27; Mismatches 52; Indels 6; Gaps 4;

QY 80 GGERADSVLSAKLEDI--ASENDWVVERMAPCLTGSQIHQITLKNDPVGGL-L 136  
 DB 38 GGDRETTMTNIDHTRVNGINNDVITVTHDAVPPILQRIKENIEVAXX-YGAVDT 95

QY 137 ALSSHDTL-KHYDGTITATIDRKHWRALTPQKFKYGMALDQRTGPNP 186  
 DB 96 VIAIDITVMSKDKXNHISIPVRENMYGQGTQPSNKLDDQSYALSSSP 146

RESULT 13  
 US-09-134-001C-3574  
 Sequence 3574, Application US/09134001C  
 Patent No. 6380370  
 GENERAL INFORMATION:  
 APPLICANT: Lynn Doucette-Stamm et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
 FILE REFERENCE: GTC-007  
 CURRENT APPLICATION NUMBER: US/09/134,001C  
 CURRENT FILING DATE: 1998-08-13  
 PRIOR APPLICATION NUMBER: US 60/064,964  
 PRIOR FILING DATE: 1997-11-08  
 PRIOR APPLICATION NUMBER: US 60/055,779  
 PRIOR FILING DATE: 1997-08-14  
 NUMBER OF SEQ ID NOS: 5674  
 SEQ ID NO 3574  
 LENGTH: 531  
 TYPE: PRT  
 ORGANISM: Staphylococcus epidermidis  
 US-09-134-001C-3574

Query Match 7.4%; Score 89; DB 4; Length 531;  
 Best Local Similarity 22.9%; Pred. No. 0.13;  
 Matches 49; Conservative 37; Mismatches 72; Indels 56; Gaps 11;

QY 29 LPLAGTIVERTITRLLESDFQKAVVAISVEDPYWPELSAKHPDITTAQGERA--- 85  
 DB 178 LPAABDPVKQELVEDIGID--QEDVVLASAKSNIGIELEKIYDVVPADGGEAPLK 235  
 QY 86 ----DS-----VLSALKALEDIASENDWVVERMAPCLTGSQI-----HIQ 123  
 DB 236 ALTFDESDPYPRKVISIRIIDGVYAGRI-----KMAWGKFEVTEVGIMPKPLP 269  
 QY 124 IDTLKNDPVGGLALSSHDTLKHYD--GDTITAT--IDRK-HVWYALTPQNF----- 170  
 DB 290 VEELTVGDVGYIIA-----SIKVVDSRVDGITITLAERPADPLOGYKKKPMVFCGLFP 344  
 QY 171 ----KYGMALDQRTGPNPVAVD--EASALELIG 199  
 DB 345 IDNKDYNDRFALEKQLNDASLEFPBESSQALG 378

RESULT 14  
 US-08-971-782-2  
 Sequence 2, Application US/08971782  
 Patent No. 6043071  
 GENERAL INFORMATION:

APPLICANT: Wallis, Nicola G.  
 APPLICANT: Shilling, Lisa K.  
 APPLICANT: Jaworski, Deborah D.  
 APPLICANT: Wang, Min  
 APPLICANT: Mooney, Jeffrey L.  
 APPLICANT: Debonuck, Christine M.  
 APPLICANT: Zhong, Yi Yi  
 TITLE OF INVENTION: No. 6043071el GIMU  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSER: Dechert Price & Rhoads  
 STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
 CITY: Philadelphia  
 STATE: PA  
 COUNTRY: US  
 ZIP: 19103

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/971,782  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/050,996  
 FILING DATE: 26-JUN-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Dickinson, Todd Q  
 REGISTRATION NUMBER: 28,354  
 REFERENCE/DOCKET NUMBER: GM10024  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 215-994-2252  
 TELEFAX: 215-994-2222

TELEX:  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 459 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-09-971-782-2

Query Match 7.4%; Score 88.5; DB 3; Length 459;  
 Best Local Similarity 23.9%; Pred. No. 0.12;  
 Matches 48; Conservative 32; Mismatches 74; Indels 47; Gaps 9;

QY 8 MAVVPAAGYGMQADRPQYPLPLAGTIVERTITRLLESDFQKAVVAISVEDPYWPEL 67  
 DB 4 FALILAGGTNRKSLPVLKHWAGISLKH-----VRSVG-ALGPE----- 46  
 QY 68 SLAKHPDITTAQGERADSVLSAKLEDIASENDWVVERMAPCLTGSQIHQI 124  
 DB 47 -----KTVVVERK--MELVBEVLACGTEFVYQSELOGTGHAVMTPEPILEGSGHTLV 98  
 QY 125 DTLKNDPVGGLALSSHDTLKHYDGTITATIDRKHWRALTPQNF--FKYGMAL-----R 176  
 DB 99 -----IAGDPPLITGESLKNL-----IDFHNNHNVATILPAEDNDFGVRGRIVRNDNA 147  
 QY 177 DALQRTGPNPVAVD--EASALEL 197  
 DB 148 EVLRNVEQDADDFEKOIKEI 168

RESULT 15  
 US-09-309-026-2  
 Sequence 2, Application US/09309026  
 Patent No. 6204042  
 GENERAL INFORMATION:  
 APPLICANT: Wallis, Nicola G.  
 APPLICANT: Shilling, Lisa K.  
 APPLICANT: Jaworski, Deborah D.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 15:54:19 : Search time 42.3225 Seconds  
(without alignments/sec  
3044.503 Million cell updates/sec

Title: US-09-941-947a-6  
Perfect score: 3216  
Sequence: 1 MKLTDYPLLNKIHFPADIR.....LSLVGLSKGLATIEQPCA 620

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PC1\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10C\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3216	100.0	620	10	US-09-934-903-2	Sequence 2, Appl1
2	3216	100.0	620	10	US-09-934-868-62	Sequence 62, Appl1
3	3216	100.0	620	11	US-09-941-947a-6	Sequence 6, Appl1
4	2068	64.7	628	12	US-10-369-493-13998	Sequence 13998, A
5	2068	64.4	627	12	US-10-381-779-15	Sequence 15, Appl1
6	2014.5	62.6	620	9	US-09-815-242-14083	Sequence 14083, A
7	2004.5	62.3	619	12	US-10-369-493-181	Sequence 481, Appl1
8	2003.5	62.3	620	12	US-10-369-493-21174	Sequence 21174, A
9	2000.5	62.2	620	9	US-09-815-242-10068	Sequence 10068, A
10	2000.5	62.2	620	12	US-10-381-779-31	Sequence 31, Appl1
11	1992.5	62.0	620	12	US-10-369-493-731	Sequence 731, Appl1
12	1992.5	62.0	623	12	US-10-369-493-15911	Sequence 15911, A
13	1992.5	62.0	623	12	US-10-369-493-16281	Sequence 16281, A
14	1992.5	62.0	653	12	US-10-369-493-15540	Sequence 15540, A
15	1990.5	61.9	619	12	US-10-369-493-7428	Sequence 7428, Appl1

16	1985.5	61.7	616	12	US-10-369-493-4669	Sequence 4669, Appl1
17	1971	61.3	615	12	US-10-369-493-9272	Sequence 9272, Appl1
18	1968.5	61.2	618	12	US-10-369-493-8507	Sequence 8507, Appl1
19	1967	61.2	613	12	US-10-369-493-9367	Sequence 9367, Appl1
20	1951	60.7	670	12	US-10-369-493-17644	Sequence 17644, A
21	1937.5	60.2	625	12	US-10-381-779-33	Sequence 33, Appl1
22	1776	55.2	637	12	US-10-381-779-32	Sequence 32, Appl1
23	1623	50.5	624	12	US-10-369-493-10188	Sequence 10188, A
24	1620.5	50.4	628	12	US-10-381-779-18	Sequence 18, Appl1
25	1620.5	50.4	628	12	US-10-369-493-61	Sequence 61, Appl1
26	1555	48.4	628	12	US-10-369-493-20596	Sequence 20596, A
27	1543.5	48.0	637	12	US-10-369-493-12290	Sequence 12290, A
28	1532.5	47.7	641	12	US-10-381-779-3	Sequence 3, Appl1
29	1530.5	47.6	629	12	US-10-369-493-10848	Sequence 10848, A
30	1523.5	47.4	640	12	US-10-369-493-16942	Sequence 16942, A
31	1503.5	46.8	654	12	US-10-369-493-7893	Sequence 7893, Appl1
32	1499.5	46.6	630	12	US-10-369-493-14775	Sequence 14775, A
33	1499.5	46.6	634	12	US-10-369-493-11753	Sequence 11753, A
34	1499.5	46.6	635	12	US-10-369-493-14237	Sequence 14237, A
35	1499.5	46.6	635	12	US-10-369-493-14928	Sequence 14928, A
36	1491.5	46.4	648	12	US-10-381-779-27	Sequence 27, Appl1
37	1485.5	46.2	720	12	US-10-427-590-26	Sequence 26, Appl1
38	1482	46.1	625	12	US-10-369-493-18033	Sequence 18033, A
39	1468	45.6	614	12	US-10-369-493-10317	Sequence 10317, A
40	1459	45.4	629	12	US-10-369-493-17357	Sequence 17357, A
41	1444	44.9	635	12	US-10-369-493-20015	Sequence 20015, A
42	1444	44.9	671	12	US-10-259-194A-414	Sequence 414, Appl1
43	1442	44.8	640	12	US-10-381-779-29	Sequence 29, Appl1
44	1442	44.8	640	12	US-10-369-493-2645	Sequence 2645, Appl1
45	1439	44.7	636	12	US-10-381-779-28	Sequence 28, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-934-903-2  
Sequence 2, Application US/09934903  
Patent No. US20020102690A1  
GENERAL INFORMATION:  
APPLICANT: Kofias, Matheos  
APPLICANT: Odou, J. Martin  
APPLICANT: Schenzle, Andreas J.  
APPLICANT: Temd, Jean-Francois  
APPLICANT: Rouverie, Pierre  
APPLICANT: Picataggio, Stephen  
TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production  
FILE REFERENCE: CL1646 US NA  
CURRENT APPLICATION NUMBER: US/09/934,903  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/229,907  
PRIOR FILING DATE: September 1, 2001  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 2  
LENGTH: 620  
TYPE: PRT  
ORGANISM: Methylobionas 16a  
FEATURE:  
OTHER INFORMATION: Amino acid sequences encoded by ORF1  
US-09-934-903-2

Query Match 100.0% Score 3216; DB 10; Length 620;  
Best Local Similarity 100.0% Pred. No. 1.1e-298; Indels 0; Gaps 0;  
Matches 620; Conservative 0; Mismatches 0;  
QY 1 MKLTDYPLLNKIHFPADIRALSKQDLADAVGVYLTHTVSISSGHPAAGLTVELTV 60  
DB 1 MKLTDYPLLNKIHFPADIRALSKQDLADAVGVYLTHTVSISSGHPAAGLTVELTV 60  
QY 61 ALHYVNTVVDQLVNVDVGHQAFPHILVGRKRMFTITLIGSVSAFPARDESYAPFGV 120

DB 61 ALHYFNTPVQDLVWDVGHQAVPEKILITGRKERMTIRITIGVSAPAPABESYDAFGVG 120  
QY 121 HSTSTISAAIGMAIASQLRGDEKXVVAIIIGDGSITGMAVEANNAHADVANNLLVILNDN 180  
DB 121 HSTSTISAAIGMAIASQLRGDEKXVVAIIIGDGSITGMAVEANNAHADVANNLLVILNDN 180  
QY 181 DMSISPPVGMANNLITVLSKFPYSSVRESKKAALAMPBWELARKTEBHVGMIVPGT 240  
DB 181 DMSISPPVGMANNLITVLSKFPYSSVRESKKAALAMPBWELARKTEBHVGMIVPGT 240  
QY 241 LPEELGFRVYFPIGHDVEMLVSTLEMLKDLTGVPFLAVYTKKGKGYAPAKDPLAYHGV 300  
DB 241 LPEELGFRVYFPIGHDVEMLVSTLEMLKDLTGVPFLAVYTKKGKGYAPAKDPLAYHGV 300  
QY 301 PAFPTKDFLKPAPSPHPITTEVFGRMLCDMAADDERLLGITPAMRBGSLVFSQKFP 360  
DB 301 PAFPTKDFLKPAPSPHPITTEVFGRMLCDMAADDERLLGITPAMRBGSLVFSQKFP 360  
QY 361 NRYFDVAIAEQHAVALTAAGACQCAKEPVVAIYSTFLORGVDOLIHVVALONTLMFLALDR 420  
DB 361 NRYFDVAIAEQHAVALTAAGACQCAKEPVVAIYSTFLORGVDOLIHVVALONTLMFLALDR 420  
QY 421 AGLVGPDGPTHAGAFDYSYKRCIPNMLIMAPADENECROMLTGFGHGGASVRYPRGKG 480  
DB 421 AGLVGPDGPTHAGAFDYSYKRCIPNMLIMAPADENECROMLTGFGHGGASVRYPRGKG 480  
QY 481 PGAAIDPPLTLALBEGKAEVRHGGSRILILANGSVTPAVAGAKQAGATVYNNKFPVKEPDQ 540  
DB 481 PGAAIDPPLTLALBEGKAEVRHGGSRILILANGSVTPAVAGAKQAGATVYNNKFPVKEPDQ 540  
QY 541 ALVLELATHTDVFTVEENVIAGAGSAINTFLOAQKVLMPVCNIGLPRFVQGSREEL 600  
DB 541 ALVLELATHTDVFTVEENVIAGAGSAINTFLOAQKVLMPVCNIGLPRFVQGSREEL 600  
QY 601 LSLVGLDSKGLATITEQFCA 620  
DB 601 LSLVGLDSKGLATITEQFCA 620

RESULT 2  
US-09-934-868-62  
; Sequence 62, Application US/09934868  
; Patent No. US20020137130A1  
; GENERAL INFORMATION:  
; APPLICANT: Koffas, Mattheos  
; APPLICANT: Odom, James M.  
; APPLICANT: Schenazle, Andreas J.  
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN  
; FILE REFERENCE: CL1596 US NA  
; CURRENT APPLICATION NUMBER: US/09/934,868  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/229,858  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 62  
; LENGTH: 620  
; TYPE: prt  
; ORGANISM: Methylobionas 16a  
; FEATURE:  
; OTHER INFORMATION: Amino acid sequences encoded by DXS  
US-09-934-868-62

Query Match 100.0%; Score 3216; DB 10; Length 620;  
Best Local Similarity 100.0%; Pred. No. 1,1e-298;  
Matches 620; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMLTTDYPLLNKNIHTPADIRALSKDQLQDLADBVAGVYTHTVSISGGHPPAGLCTVELTY 60  
DB 1 KMLTTDYPLLNKNIHTPADIRALSKDQLQDLADBVAGVYTHTVSISGGHPPAGLCTVELTY 60  
QY 61 ALHYFNTPVQDLVWDVGHQAVPEKILITGRKERMTIRITIGVSAPAPABESYDAFGVG 120

DB 61 ALHYFNTPVQDLVWDVGHQAVPEKILITGRKERMTIRITIGVSAPAPABESYDAFGVG 120  
QY 121 HSTSTISAAIGMAIASQLRGDEKXVVAIIIGDGSITGMAVEANNAHADVANNLLVILNDN 180  
DB 121 HSTSTISAAIGMAIASQLRGDEKXVVAIIIGDGSITGMAVEANNAHADVANNLLVILNDN 180  
QY 181 DMSISPPVGMANNLITVLSKFPYSSVRESKKAALAMPBWELARKTEBHVGMIVPGT 240  
DB 181 DMSISPPVGMANNLITVLSKFPYSSVRESKKAALAMPBWELARKTEBHVGMIVPGT 240  
QY 241 LPEELGFRVYFPIGHDVEMLVSTLEMLKDLTGVPFLAVYTKKGKGYAPAKDPLAYHGV 300  
DB 241 LPEELGFRVYFPIGHDVEMLVSTLEMLKDLTGVPFLAVYTKKGKGYAPAKDPLAYHGV 300  
QY 301 PAFPTKDFLKPAPSPHPITTEVFGRMLCDMAADDERLLGITPAMRBGSLVFSQKFP 360  
DB 301 PAFPTKDFLKPAPSPHPITTEVFGRMLCDMAADDERLLGITPAMRBGSLVFSQKFP 360  
QY 361 NRYFDVAIAEQHAVALTAAGACQCAKEPVVAIYSTFLORGVDOLIHVVALONTLMFLALDR 420  
DB 361 NRYFDVAIAEQHAVALTAAGACQCAKEPVVAIYSTFLORGVDOLIHVVALONTLMFLALDR 420  
QY 421 AGLVGPDGPTHAGAFDYSYKRCIPNMLIMAPADENECROMLTGFGHGGASVRYPRGKG 480  
DB 421 AGLVGPDGPTHAGAFDYSYKRCIPNMLIMAPADENECROMLTGFGHGGASVRYPRGKG 480  
QY 481 PGAAIDPPLTLALBEGKAEVRHGGSRILILANGSVTPAVAGAKQAGATVYNNKFPVKEPDQ 540  
DB 481 PGAAIDPPLTLALBEGKAEVRHGGSRILILANGSVTPAVAGAKQAGATVYNNKFPVKEPDQ 540  
QY 541 ALVLELATHTDVFTVEENVIAGAGSAINTFLOAQKVLMPVCNIGLPRFVQGSREEL 600  
DB 541 ALVLELATHTDVFTVEENVIAGAGSAINTFLOAQKVLMPVCNIGLPRFVQGSREEL 600  
QY 601 LSLVGLDSKGLATITEQFCA 620  
DB 601 LSLVGLDSKGLATITEQFCA 620

RESULT 3  
US-09-941-947a-6  
; Sequence 6, Application US/09941947A  
; Publication No. US20030003528A1  
; GENERAL INFORMATION:  
; APPLICANT: Brzostowicz, Patricia C.  
; APPLICANT: Cheng, Qiong  
; APPLICANT: DiCosimo, Deana J.  
; APPLICANT: Koffas, Mattheos  
; APPLICANT: Miller, Edward S. Jr.  
; APPLICANT: Odom, J. Martin  
; APPLICANT: Picataggio, Steve  
; APPLICANT: Rouviere, Pierre E.  
; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE  
; FILE REFERENCE: CL1903 US NA  
; CURRENT APPLICATION NUMBER: US/09/941,947A  
; PRIOR FILING DATE: 2001-09-01  
; PRIOR APPLICATION NUMBER: 60/229,907  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,858  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 6  
; LENGTH: 620  
; TYPE: prt  
; ORGANISM: Methylobionas 16a  
US-09-941-947a-6

Query Match 100.0%; Score 3216; DB 11; Length 620;  
Best Local Similarity 100.0%; Pred. No. 1,1e-298;  
Matches 620; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MKLTTDYPPLKNIHTPADIRALSKDQLOQADAEVGYLTHTVSISGGHFAAGLTVETLV 60
DB 1 MKLTTDYPPLKNIHTPADIRALSKDQLOQADAEVGYLTHTVSISGGHFAAGLTVETLV 60
QY 61 ALHYFNTFVDQVWDVGHQAVPHKILNCRKEMPTIRTLGVSAFPARDESEYDARVG 120
DB 61 ALHYFNTFVDQVWDVGHQAVPHKILNCRKEMPTIRTLGVSAFPARDESEYDARVG 120
QY 121 HSTSTISALGMAIASQLEGEDEKKNVAIIGDSTIGMAAYEAMNAGVANNLVIINDN 180
DB 121 HSTSTISALGMAIASQLEGEDEKKNVAIIGDSTIGMAAYEAMNAGVANNLVIINDN 180
QY 121 HSTSTISALGMAIASQLEGEDEKKNVAIIGDSTIGMAAYEAMNAGVANNLVIINDN 180
DB 121 HSTSTISALGMAIASQLEGEDEKKNVAIIGDSTIGMAAYEAMNAGVANNLVIINDN 180
QY 181 DMSISPPVGMANNVLTQVLSKRYSSVREESKKALAKMPSWELARKTEEHVKMIVGT 240
DB 181 DMSISPPVGMANNVLTQVLSKRYSSVREESKKALAKMPSWELARKTEEHVKMIVGT 240
QY 241 LFEELGPNYFGPIIDGHVEMVSTLENKDLTPVPLHVTYKKGKGYAPAKDPLAHGV 300
DB 241 LFEELGPNYFGPIIDGHVEMVSTLENKDLTPVPLHVTYKKGKGYAPAKDPLAHGV 300
QY 301 PAFDPTKDFLPRAADSPHYTYEVFGNMLCDMAADERLLGITPAMEGSGLVEFSQKFP 360
DB 301 PAFDPTKDFLPRAADSPHYTYEVFGNMLCDMAADERLLGITPAMEGSGLVEFSQKFP 360
QY 361 NRYFVAIAEQAHTVLAAGAACQAKPVVAIYSTFLQGVQDLIHVALQNLDMLEFALDR 420
DB 361 NRYFVAIAEQAHTVLAAGAACQAKPVVAIYSTFLQGVQDLIHVALQNLDMLEFALDR 420
QY 421 AGVGPDPGPTAGADPVSYMRCIPNMLIMAPADENECOMLTGPOHNGPASYRYPGK 480
DB 421 AGVGPDPGPTAGADPVSYMRCIPNMLIMAPADENECOMLTGPOHNGPASYRYPGK 480
QY 481 PGAAIDPFLTLALIEIGKAEVRHGSRIALIAWGSMTTPAVEAGKOLGATVNMREVPKDFQ 540
DB 481 PGAAIDPFLTLALIEIGKAEVRHGSRIALIAWGSMTTPAVEAGKOLGATVNMREVPKDFQ 540
QY 541 ALVLELARTHDFVFEVENVIAAGGSAINTFLQAKVLMPCVNICIGLDPFVQSGSEEL 600
DB 541 ALVLELARTHDFVFEVENVIAAGGSAINTFLQAKVLMPCVNICIGLDPFVQSGSEEL 600
QY 601 LSLVGLDSKGLIATIEQFCA 620
DB 601 LSLVGLDSKGLIATIEQFCA 620

RESULT 4
US-10-369-493-13998
; Sequence 13998, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiandeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13998
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Pseudomonas fluorescens
US-10-369-493-13998

Query Match 64.7%; Score 2080; DB 12; Length 628;
Best Local Similarity 63.2%; Pred. No. 7e-190;
Matches 387; Conservative 101; Mismatches 120; Indels 4; Gaps 2;

```

```

QY 8 PLIKNIHTPADIRALSKDQLOQADAEVGYLTHTVSISGGHFAAGLTVETLV 67
DB 16 PLIKNIHTPADIRALSKDQLOQADAEVGYLTHTVSISGGHFAAGLTVETLV 67
QY 68 TPVDQVWDVGHQAVPHKILNCRKEMPTIRTLGVSAFPARDESEYDARVG 127
DB 76 TPVDQVWDVGHQAVPHKILNCRKEMPTIRTLGVSAFPARDESEYDARVG 127
QY 128 AALGMAIASQLEGEDEKKNVAIIGDSTIGMAAYEAMNAGVANNLVIINDN 187
DB 136 AALGMAIASQLEGEDEKKNVAIIGDSTIGMAAYEAMNAGVANNLVIINDN 187
QY 188 VGMANNVLTQVLSKRYSSVREESKKALAKMPSWELARKTEEHVKMIVGT 247
DB 196 VGMANNVLTQVLSKRYSSVREESKKALAKMPSWELARKTEEHVKMIVGT 247
QY 248 NRYFVAIAEQAHTVLAAGAACQAKPVVAIYSTFLQGVQDLIHVALQNLDMLEFALDR 307
DB 256 NRYFVAIAEQAHTVLAAGAACQAKPVVAIYSTFLQGVQDLIHVALQNLDMLEFALDR 307
QY 308 DFLPRAAP--SHPHYTYEVFGNMLCDMAADERLLGITPAMEGSGLVEFSQKFP 365
DB 316 --APAAAPKAGGPKYSGVFGNMLCDMAADERLLGITPAMEGSGLVEFSQKFP 365
QY 366 VALAQAHTVLAAGAACQAKPVVAIYSTFLQGVQDLIHVALQNLDMLEFALDR 425
DB 374 VALAQAHTVLAAGAACQAKPVVAIYSTFLQGVQDLIHVALQNLDMLEFALDR 425
QY 426 PGDPPTAGADPVSYMRCIPNMLIMAPADENECOMLTGPOHNGPASYRYPGK 485
DB 434 PGDPPTAGADPVSYMRCIPNMLIMAPADENECOMLTGPOHNGPASYRYPGK 485
QY 486 DPTLALBIGAEVRHGSRIALIAWGSMTTPAVEAGKOLGATVNMREVPKDFQ 545
DB 494 DPTLALBIGAEVRHGSRIALIAWGSMTTPAVEAGKOLGATVNMREVPKDFQ 545
QY 546 LARTHDVPTVSENVIAAGGSAINTFLQAKVLMPCVNICIGLDPFVQSGSEEL 605
DB 554 LARTHDVPTVSENVIAAGGSAINTFLQAKVLMPCVNICIGLDPFVQSGSEEL 605
QY 606 LSLVGLDSKGLIATIEQ 617
DB 614 LSLVGLDSKGLIATIEQ 617

RESULT 5
US-10-381-779-15
; Sequence 15, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jessem, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12304/002051
; CURRENT APPLICATION NUMBER: US/10/381,779
; PRIOR FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-381-779-15

Query Match 64.4%; Score 2069.5; DB 12; Length 627;
Best Local Similarity 63.4%; Pred. No. 7.1e-189;
Matches 388; Conservative 100; Mismatches 117; Indels 7; Gaps 2;

```

QY 8 PLKNIHTPADIRALSKDQLOQLADVRGYLTHTVSISGGHFAAGISTVELTVALHYVN 67  
 DB 16 PLLRASSPAPLRLGENDLETTADELRQYLYLVGGTGHFGAGLGVLTALHYVD 75  
 QY 68 TPVDQVMDVGHQAVPHKILTRKREMPRTIRTTGGVAPAPADESEYDAVGVSSTIS 127  
 DB 76 TPDRQVMDVGHQAVPHKILTRERELMGTLAGKGLAAPPRAESEYDITGVGHSSTIS 135  
 QY 128 AALGMATASQIRGEDKXKVALIGDSITGGMAYANMHAQDVANLLVILINDMSISPP 187  
 DB 136 AALGMATASQIRGEDKXKVALIGDSITGGMAYANMHAQDVANLLVILINDMSISPP 195  
 QY 188 VGMANNVLTLYLSSKFFSVSVESKALAKMPVMEIARKEEHEVGMIVPGLTFEELG 247  
 DB 196 VGMANNVLTLYLSSKFFSVSVESKALAKMPVMEIARKEEHEVGMIVPGLTFEELG 255  
 QY 248 NYFGPIDGHVEMLVSTLENIKDITGVPFLHVTVTKKGKGYAPAKOPLAYHGVAPDPK 307  
 DB 256 NYFGPIDGHVEMLVSTLENIKDITGVPFLHVTVTKKGKGYAPAKOPLAYHGVAPDPK 310  
 QY 308 DELPKAAE--SEHPTYTEVFGRWICDMAAOBERLLGITPAMRESGGLVERGKPPNRYED 365  
 DB 311 LEAPSAKXKTGGKPTSSVFCQMLCDMAADQARLLGITPAMRESGGLVERGKPPNRYED 370  
 QY 366 VALAEOHAVTLAAGQACQAGAPVVAIYSTFLQRTDQILHVALQNLMDLFDLDRAGLVG 425  
 DB 371 VALAEOHAVTLAAGQACQAGAPVVAIYSTFLQRTDQILHVALQNLMDLFDLDRAGLVG 430  
 QY 426 PDGPTHAGAPDYSYKRCIPNMLIAPADNECROMLTGROHGPASVRRPRKGPAAI 465  
 DB 431 EDGPTHAGAPDYSYKRCIPNMLIAPADNECROMLTGROHGPASVRRPRKGPAAI 490  
 QY 466 DPTLTALRIGKAEVYHSGRIATILAMGSMVTPAVEAGKOLGATVNNRFPVCPDQALVE 545  
 DB 491 DPTLTALRIGKAEVYHSGRIATILAMGSMVTPAVEAGKOLGATVNNRFPVCPDQALVE 550  
 QY 546 IARTHDVFTVYENVIAGAGSANTFLOAQKVLMPVCNIGLPRFPYQSGREBELSLVG 605  
 DB 551 IARTHDVFTVYENVIAGAGSANTFLOAQKVLMPVCNIGLPRFPYQSGREBELSLVG 610  
 QY 606 LDSKGLIATIBQ 617  
 DB 611 LDSKGLIATIBQ 622

RESULT 6  
 US-09-815-242-14083  
 ; Sequence 14083, Application US/09815242  
 ; Patent No. US20020061569A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Karl L.  
 ; APPLICANT: Zyskind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.  
 ; APPLICANT: Xu, H. Howard  
 ; TITLE OF INVENTION: Identification of Essential Genes in  
 ; FILE REFERENCE: ELITRA.011A  
 ; CURRENT APPLICATION NUMBER: US/09/815,242  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27

QY 7 YPLKNIHTPADIRALSKDQLOQLADVRGYLTHTVSISGGHFAAGISTVELTVALHYVN 66  
 DB 8 YPLKNIHTPADIRALSKDQLOQLADVRGYLTHTVSISGGHFAAGISTVELTVALHYVN 67  
 QY 67 NTPVDQVMDVGHQAVPHKILTRKREMPRTIRTTGGVAPAPADESEYDAVGVSSTIS 126  
 DB 68 NTPVDQVMDVGHQAVPHKILTRKREMPRTIRTTGGVAPAPADESEYDAVGVSSTIS 127  
 QY 127 SAALGMATASQIRGEDKXKVALIGDSITGGMAYANMHAQDVANLLVILINDMSISPP 186  
 DB 128 SAALGMATASQIRGEDKXKVALIGDSITGGMAYANMHAQDVANLLVILINDMSISPP 187  
 QY 187 PVGMANNVLTLYLSSKFFSVSVESKALAKMPVMEIARKEEHEVGMIVPGLTFEELG 246  
 DB 188 PVGMANNVLTLYLSSKFFSVSVESKALAKMPVMEIARKEEHEVGMIVPGLTFEELG 247  
 QY 247 FNYFGPIDGHVEMLVSTLENIKDITGVPFLHVTVTKKGKGYAPAKOPLAYHGVAPDPK 306  
 DB 248 FNYFGPIDGHVEMLVSTLENIKDITGVPFLHVTVTKKGKGYAPAKOPLAYHGVAPDPK 307  
 QY 307 KDFLPKAAE--SEHPTYTEVFGRWICDMAAOBERLLGITPAMRESGGLVERGKPPNRYED 366  
 DB 308 KDFLPKAAE--SEHPTYTEVFGRWICDMAAOBERLLGITPAMRESGGLVERGKPPNRYED 367  
 QY 367 VALAEOHAVTLAAGQACQAGAPVVAIYSTFLQRTDQILHVALQNLMDLFDLDRAGLVG 426  
 DB 368 VALAEOHAVTLAAGQACQAGAPVVAIYSTFLQRTDQILHVALQNLMDLFDLDRAGLVG 427  
 QY 427 DPTLTALRIGKAEVYHSGRIATILAMGSMVTPAVEAGKOLGATVNNRFPVCPDQALVE 545  
 DB 428 DPTLTALRIGKAEVYHSGRIATILAMGSMVTPAVEAGKOLGATVNNRFPVCPDQALVE 546  
 QY 546 IARTHDVFTVYENVIAGAGSANTFLOAQKVLMPVCNIGLPRFPYQSGREBELSLVG 605  
 DB 547 IARTHDVFTVYENVIAGAGSANTFLOAQKVLMPVCNIGLPRFPYQSGREBELSLVG 606  
 QY 606 LDSKGLIATIBQ 617  
 DB 607 LDSKGLIATIBQ 618

RESULT 7  
 US-10-369-493-481  
 ; Sequence 481, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; PLANTS WITH IMPROVED PROPERTIES

QY 7 YPLKNIHTPADIRALSKDQLOQLADVRGYLTHTVSISGGHFAAGISTVELTVALHYVN 66  
 DB 8 YPLKNIHTPADIRALSKDQLOQLADVRGYLTHTVSISGGHFAAGISTVELTVALHYVN 67  
 QY 67 NTPVDQVMDVGHQAVPHKILTRKREMPRTIRTTGGVAPAPADESEYDAVGVSSTIS 126  
 DB 68 NTPVDQVMDVGHQAVPHKILTRKREMPRTIRTTGGVAPAPADESEYDAVGVSSTIS 127  
 QY 127 SAALGMATASQIRGEDKXKVALIGDSITGGMAYANMHAQDVANLLVILINDMSISPP 186  
 DB 128 SAALGMATASQIRGEDKXKVALIGDSITGGMAYANMHAQDVANLLVILINDMSISPP 187  
 QY 187 PVGMANNVLTLYLSSKFFSVSVESKALAKMPVMEIARKEEHEVGMIVPGLTFEELG 246  
 DB 188 PVGMANNVLTLYLSSKFFSVSVESKALAKMPVMEIARKEEHEVGMIVPGLTFEELG 247  
 QY 247 FNYFGPIDGHVEMLVSTLENIKDITGVPFLHVTVTKKGKGYAPAKOPLAYHGVAPDPK 306  
 DB 248 FNYFGPIDGHVEMLVSTLENIKDITGVPFLHVTVTKKGKGYAPAKOPLAYHGVAPDPK 307  
 QY 307 KDFLPKAAE--SEHPTYTEVFGRWICDMAAOBERLLGITPAMRESGGLVERGKPPNRYED 366  
 DB 308 KDFLPKAAE--SEHPTYTEVFGRWICDMAAOBERLLGITPAMRESGGLVERGKPPNRYED 367  
 QY 367 VALAEOHAVTLAAGQACQAGAPVVAIYSTFLQRTDQILHVALQNLMDLFDLDRAGLVG 426  
 DB 368 VALAEOHAVTLAAGQACQAGAPVVAIYSTFLQRTDQILHVALQNLMDLFDLDRAGLVG 427  
 QY 427 DPTLTALRIGKAEVYHSGRIATILAMGSMVTPAVEAGKOLGATVNNRFPVCPDQALVE 545  
 DB 428 DPTLTALRIGKAEVYHSGRIATILAMGSMVTPAVEAGKOLGATVNNRFPVCPDQALVE 546  
 QY 546 IARTHDVFTVYENVIAGAGSANTFLOAQKVLMPVCNIGLPRFPYQSGREBELSLVG 605  
 DB 547 IARTHDVFTVYENVIAGAGSANTFLOAQKVLMPVCNIGLPRFPYQSGREBELSLVG 606  
 QY 606 LDSKGLIATIBQ 617  
 DB 607 LDSKGLIATIBQ 618

```

; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 481
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-481

```

```

Query Match      62.3%; Score 2004.5; DB 12; Length 619;
Best Local Similarity 61.8%; Pred. No. 1.2e-182;
Matches 378; Conservative 102; Mismatches 129; Indels 3; Gaps 3;

```

```

QY 7 YPLKNHHTPADIRALSKDQLQGLADEVRGYLTHTVSISGGHFAAGLGTVELTVALHYVF 66
DB 8 YPTLALATPEBELRLPDTLPKLCDELKQPLNVSRSCHFAAGLGTVELTVALHYVF 67
QY 67 NTPVDQLVWDVGHQAYPHKILTRKREMPTRITLGGVSAFPARDESEYDAFGVGHSTSI 126
DB 68 KTFPDMVLVWDVGHQAYPHKILTRGRDRIINTIRKQKGLHPFWRDESEYDTLCVGHSTSI 127
QY 127 SAALGKALASQLRGDKKXVALIGDGSITGGAAAYAMNHAGDVANLLVILINDMSTSP 186
DB 128 SAGLGKALAAKEDKGRKTVCVIGGALTAGAFAAMNHAGDIDPDMVLINDMSTSP 187
QY 187 PYGAMNNVLTFLYSSKFSYVSEBSKALAKPVSWEELARTBEHVKMIVPGLTFEELG 246
DB 188 NVGALNNHLAHLSSKLTTLREGSKYFNSLPPIKELKTEHLMKMMVPGTLFEEIG 247
QY 247 FNYFGPIDGHVEMLVSTLENKDI:GCVPLHVTYKKGKGYAPAKDPLAYHGVAPDPT 306
DB 248 FNYIPVGDHVDLALTQTLKNMRBELKGPQFLHMTKKRGYAPAKDPIGWHAVPKEDPS 307
QY 307 KDFLKAAPSPHPTTEYVGRMLCPMAADERLIGTPAMEGSGLVESQKPFNRKYDV 366
DB 308 TGSLSKSSPT-RPTTSKIPGWLCEBAHDKKMAITPAMEGSGMRFSEYEQYFDV 366
QY 367 ALAEQHAVTLAAGACQCAKPVVAIYSTFLQRYDQLIHDVALQNLDMFLALDRAGLVGP 426
DB 367 ALAEQHAVTFPAAGLIGCYKEPVVAIYSTFLQRAVDQLIHDVAIQNLPLFAIDGGLVGA 426
QY 427 DGPHTAGAFDYSVMEICINMLIMAPADNECRQMLTGFGCH--HBPASVVRPGRGPAAI 485
DB 427 DQHTQGAFDLSFLCINMVMAPSDNECRQMLHTGYQYQSGPVAVRIPRAGTGAEL 486
QY 486 DPTLALBIEGKAEVHHGSRITAILAMGSMVTPAVEAGKOLGATVVMRFPVCPDQALVE 545
DB 487 GP-LEILPIGKGVIRQGEKIALINFGTLLENALQAAESLNAVTVDRFPVLPDKEIVLE 545
QY 546 LARTHDVPTVBEENVLAGAGSAINFLQAKVILMPVYNIGLPRFPBGSSREELSLV 605
DB 546 MASRIDLVLEENAVIMGAGSGVVELLMQTNQSVSLNIGLPYFLPQGTQAEHSDIG 605
QY 606 LQSGKGLIATIEQ 617
DB 606 LQSGKGLIATIEQ 617

```

```

RESULT 8
US-10-369-493-21174
; Sequence 21174, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

```

```

; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21174
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-21174

```

```

Query Match      62.3%; Score 2003.5; DB 12; Length 620;
Best Local Similarity 61.1%; Pred. No. 1.5e-182;
Matches 377; Conservative 102; Mismatches 131; Indels 7; Gaps 4;

```

```

QY 7 YPLKNHHTPADIRALSKDQLQGLADEVRGYLTHTVSISGGHFAAGLGTVELTVALHYVF 66
DB 8 YPTLALATPEBELRLPDTLPKLCDELKQPLNVSRSCHFAAGLGTVELTVALHYVF 67
QY 67 NTPVDQLVWDVGHQAYPHKILTRKREMPTRITLGGVSAFPARDESEYDAFGVGHSTSI 126
DB 68 KTFPDMVLVWDVGHQAYPHKILTRGRDRIINTIRKQKGLHPFWRDESEYDTLCVGHSTSI 127
QY 127 SAALGKALASQLRGDKKXVALIGDGSITGGAAAYAMNHAGDVANLLVILINDMSTSP 186
DB 128 SAGLGKALAAKEDKGRKTVCVIGGALTAGAFAAMNHAGDIDPDMVLINDMSTSP 187
QY 187 PYGAMNNVLTFLYSSKFSYVSEBSKALAKPVSWEELARTBEHVKMIVPGLTFEELG 246
DB 188 NVGALNNHLAHLSSKLTTLREGSKYFNSLPPIKELKTEHLMKMMVPGTLFEEIG 247
QY 247 FNYFGPIDGHVEMLVSTLENKDI:GCVPLHVTYKKGKGYAPAKDPLAYHGVAPDPT 306
DB 248 FNYIPVGDHVDLALTQTLKNMRBELKGPQFLHMTKKRGYAPAKDPIGWHAVPKEDPS 307
QY 307 KDFLKAAPSPHPTTEYVGRMLCPMAADERLIGTPAMEGSGLVESQKPFNRKYDV 366
DB 308 TGSLSKSSPT-RPTTSKIPGWLCEBAHDKKMAITPAMEGSGMRFSEYEQYFDV 366
QY 367 ALAEQHAVTLAAGACQCAKPVVAIYSTFLQRYDQLIHDVALQNLDMFLALDRAGLVGP 426
DB 367 ALAEQHAVTFPAAGLIGCYKEPVVAIYSTFLQRAVDQLIHDVAIQNLPLFAIDGGLVGA 426
QY 427 DGPHTAGAFDYSVMEICINMLIMAPADNECRQMLTGFGCH--HBPASVVRPGRGPAAI 485
DB 427 DQHTQGAFDLSFLCINMVMAPSDNECRQMLHTG--HHYQSGPVVAVYPRGTGGA 484
QY 484 AIDPTLALBIEGKAEVHHGSRITAILAMGSMVTPAVEAGKOLGATVVMRFPVCPDQALV 543
DB 485 BLQF-LEALPIGKGVIRQGEKIALINFGTLLENALQAAESLNAVTVDRFPVLPDKEIVLE 543
QY 544 LELATHDVPTVBEENVLAGAGSAINFLQAKVILMPVYNIGLPRFPBGSSREELSLV 603
DB 544 LELASHDMLVLEENAVIMGAGSGVVELLMQTNQSVSLNIGLPYFLPQGTQAEHSDIG 603
QY 604 VQIDSGKGLIATIEQFCA 620
DB 604 LGLDARGIOMAIKTLA 620

```

```

RESULT 9
US-09-815-242-10068
; Sequence 10068, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

```



APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10068  
LENGTH: 620  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-09-815-242-10068

Query Match 62.2%; Score 2000.5; DB 9; Length 620;  
Best Local Similarity 60.2%; Pred. No. 2.8e-182;  
Matches 370; Conservative 114; Mismatches 128; Indels 3; Gaps 3;  
7 YELKNHTPADIRALSDQQLADVRGLTHVTSISGHPAAGTVELTVALHYF 66  
8 YETLALVSTOELRALPESLPEKCDERRLDSVSSSHFASGLGTVELTVALHYV 67  
67 NTPVQOLVWDVGHQAYPKHILTGKREKAPTRITLGVSAFAPADESEYDAFGVSHSTSI 126  
68 NTPVQOLVWDVGHQAYPKHILTGKREKAPTRITLGVSAFAPADESEYDAFGVSHSTSI 127  
127 SATGMAINSQLRGDEKCKVAIIIGDSITGMAVEANHADVNAANLITINDMSISP 186  
128 SAGIGIYAAAEKGNRRVCYIGGATAGMAFEANHADIDPMLVILINDMSISB 187  
187 PYGAMNNYITKYLSKFTYSVREESKALAMPVWEIARKEEYKMTYPTGLFEEIG 246  
188 NVGALNNHLAQLSGLYSLSREGKCVSGVPPIKELIKRTEHIKGMVPGTLFEEIG 247  
247 FNYGFIIDGHVEMLVSTLENLKDITGFVFLHVTKKGYAPAEKDPYAHGVAPDPPT 306  
248 FNYGFIIDGHVEMLVSTLENLKDITGFVFLHVTKKGYAPAEKDPYAHGVAPDPPT 307  
307 KOTLPKAPSPHPTTEYVGRWLCPMAADDERLLGITPAMEGSGLVFESQKFPNRYFDV 366  
308 SGLPEKRS-CGLPSYSKIFGDMLCETAAKONKLAITPAMEGSGWVFSKFPDRYFDV 366  
367 AIAEQHAAVTLAAGAACQAKKPVVAIYSTFLQGVQOLIHVALQNLMLPALDRAGLVGP 426  
367 AIAEQHAAVTLAAGAACQAKKPVVAIYSTFLQGVQOLIHVALQNLMLPALDRAGLVGP 426  
427 DQPTAGADVSYMRCIPNMLIMAPADENECQMLTTPQNH-GPASYRYRGKGPAAI 485  
427 DQPTAGADVSYMRCIPNMLIMAPADENECQMLTTPQNH-GPASYRYRGKGPAAI 486  
486 DPTLTALBICAEVREHSGRIAILMAGSVTPPAVAGKQIATVNMRFVYPPDQALYLE 545  
487 TP-LEKLPKIGKIVRSGKLAIFNGTILMPAPAAVASTINATLVDMFVPLDBALILE 545  
546 LARTHDVFTVEENYIAGAGSAINFTLOAQKVLAMPVCNIGLPDRFVQSGREELLSVIG 605  
546 MASHBALVTVENALMGAGSVNIVMAHKKPVPVINTIGLPDFIPIQGTQOEWRAELG 605  
606 LDSKGLIATIBQFCA 620

Db 606 LDAAGMEKIRAMLA 620  
RESULT 10  
US-10-381-779-31  
Sequence 31, Application US/10381779  
Publication No. US20030219798A1  
GENERAL INFORMATION:  
APPLICANT: Jensen, Ravi R  
APPLICANT: Zidwick, Mary Jo  
TITLE OF INVENTION: Isoprenoid Production  
FILE REFERENCE: 12904/0020S1  
CURRENT APPLICATION NUMBER: US/10/381,779  
CURRENT FILING DATE: 2003-03-28  
PRIOR APPLICATION NUMBER: PCT/US01/30328  
PRIOR FILING DATE: 2001-09-28  
PRIOR APPLICATION NUMBER: 60/236,580  
PRIOR FILING DATE: 2000-09-29  
NUMBER OF SEQ ID NOS: 190  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 31  
LENGTH: 620  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-10-381-779-31

Query Match 62.2%; Score 2000.5; DB 12; Length 620;  
Best Local Similarity 60.2%; Pred. No. 2.8e-182;  
Matches 370; Conservative 114; Mismatches 128; Indels 3; Gaps 3;  
7 YELKNHTPADIRALSDQQLADVRGLTHVTSISGHPAAGTVELTVALHYF 66  
8 YETLALVSTOELRALPESLPEKCDERRLDSVSSSHFASGLGTVELTVALHYV 67  
67 NTPVQOLVWDVGHQAYPKHILTGKREKAPTRITLGVSAFAPADESEYDAFGVSHSTSI 126  
68 NTPVQOLVWDVGHQAYPKHILTGKREKAPTRITLGVSAFAPADESEYDAFGVSHSTSI 127  
127 SATGMAINSQLRGDEKCKVAIIIGDSITGMAVEANHADVNAANLITINDMSISP 186  
128 SAGIGIYAAAEKGNRRVCYIGGATAGMAFEANHADIDPMLVILINDMSISB 187  
187 PYGAMNNYITKYLSKFTYSVREESKALAMPVWEIARKEEYKMTYPTGLFEEIG 246  
188 NVGALNNHLAQLSGLYSLSREGKCVSGVPPIKELIKRTEHIKGMVPGTLFEEIG 247  
247 FNYGFIIDGHVEMLVSTLENLKDITGFVFLHVTKKGYAPAEKDPYAHGVAPDPPT 306  
248 FNYGFIIDGHVEMLVSTLENLKDITGFVFLHVTKKGYAPAEKDPYAHGVAPDPPT 307  
307 KOTLPKAPSPHPTTEYVGRWLCPMAADDERLLGITPAMEGSGLVFESQKFPNRYFDV 366  
308 SGLPEKRS-CGLPSYSKIFGDMLCETAAKONKLAITPAMEGSGWVFSKFPDRYFDV 366  
367 AIAEQHAAVTLAAGAACQAKKPVVAIYSTFLQGVQOLIHVALQNLMLPALDRAGLVGP 426  
367 AIAEQHAAVTLAAGAACQAKKPVVAIYSTFLQGVQOLIHVALQNLMLPALDRAGLVGP 426  
427 DQPTAGADVSYMRCIPNMLIMAPADENECQMLTTPQNH-GPASYRYRGKGPAAI 485  
427 DQPTAGADVSYMRCIPNMLIMAPADENECQMLTTPQNH-GPASYRYRGKGPAAI 486  
486 DPTLTALBICAEVREHSGRIAILMAGSVTPPAVAGKQIATVNMRFVYPPDQALYLE 545  
487 TP-LEKLPKIGKIVRSGKLAIFNGTILMPAPAAVASTINATLVDMFVPLDBALILE 545  
546 LARTHDVFTVEENYIAGAGSAINFTLOAQKVLAMPVCNIGLPDRFVQSGREELLSVIG 605  
546 MASHBALVTVENALMGAGSVNIVMAHKKPVPVINTIGLPDFIPIQGTQOEWRAELG 605  
606 LDSKGLIATIBQFCA 620

Db 606 LBAAGMEAKIKAWILA 620

## RESULT 11

US-10-369-493-731  
 ; Sequence 731, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 731  
 ; LENGTH: 620  
 ; TYPE: PRT  
 ; ORGANISM: Escherichia coli  
 ; US-10-369-493-731

Query Match 62.2%; Score 2000.5; DB 12; Length 620;  
 Best Local Similarity 60.2%; Pred. No. 2.8e-182;  
 Matches 370; Conservative 114; Mismatches 128; Indels 3; Gaps 3;

Qy 7 YPLLNKIHPTADIRALSKOQLQDLADVRGILTHYTSISGGHFAAGLGVETLVAAH 66  
 Db 8 YPLALVDSIOEIRLIPKESLPFLCDELRRYLDVSRSKSGHFAAGLGVETLVAAH 67  
 Qy 67 NTPVDOLVMDVGHQAVPHKILNCRKEMPTIRTLGVSAPFPADESEYDAFGVGHSS 126  
 Db 68 NTPFDOLVMDVGHQAVPHKILNCRKEMPTIRTLGVSAPFPADESEYDAFGVGHSS 127  
 Qy 127 SAALGMAIASQLGEGDKKVAITIGDGSITGGMAVEANNAAGDVANLVLINDMSISP 186  
 Db 128 SAQIGIAVAEKEGKRRVTCVIGDAITPAGMAFEAMNAGDIRPDMVILINDMSISB 187  
 Qy 187 PVGAMNNYLTLYVSSKFSVRESKKAALAKMPSVWEILARKTEEHVKGIVTGLFEELG 246  
 Db 188 NVGALNNHIAQLISGLYSLSREKGVSGVPIVELKRTSEHKGAVVFGTLFEELG 247  
 Qy 247 FNYEGPIDGHDVEMVSTLENLKDILTGPFVLHVTKKGQVAPAEKDPVAVGPAFDP 306  
 Db 248 FNYIGPVDGHDVGLITLTKNNRDLKGPQLHMTKKGQVAPAEKDPITFAVVPKDP 307  
 Qy 307 KDLPLRAABSPHTTYEVGKMLCDMAADDERLGLTPMREGSGLVESQTFPRRYDV 366  
 Db 308 SGCLPKRS -GGLSYSKIFQDMICETPAADNKLMAITPAMREGSGVVESSRFPORYDV 366  
 Qy 367 AINAEQAVTLAAGQAQOGAKPVVAISTFLQSGYDQLHDVALQNLMDLFDLDRAGLVP 426  
 Db 367 AINAEQAVTLAAGQAQOGAKPVVAISTFLQSGYDQLHDVALQNLMDLFDLDRAGLVP 426  
 Qy 427 DGTTHAGADPYSIMRCIPNMLMAADENECQMLTTGQHH -GPAASVRYPPKGPAAI 485  
 Db 427 DGTTHAGADPYSIMRCIPNMLMAADENECQMLTTGQHH -GPAASVRYPPKGPAAI 486  
 Qy 486 DPLTALIEIGKAEVRHSGRIALANGSWYTPAVBAGKOLGATVVMREVPKFPDAILB 545  
 Db 487 TP-LKCLPIKGIKVGKRGKALINFTGLMPPAKVABSLANTLVDMKFPVRLDAILB 545  
 Qy 546 LANTEDVFTVEENVVIAAGAGSAINTFLOAQKVLMPVCNIGLPDRFVBOGSREELSLVG 605  
 Db 546 MAASHALVTVENALIMGAGSGVNEVILMAHRAKPVVNLIGLPDFFIPIQITQEEWRAELG 605  
 Qy 606 LBSKGLIATIEQCA 620

Db 606 LBAAGMEAKIKAWILA 620

## RESULT 12

US-10-369-493-15911  
 ; Sequence 15911, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 15911  
 ; LENGTH: 623  
 ; TYPE: PRT  
 ; ORGANISM: Xanthomonas campestris  
 ; US-10-369-493-15911

Query Match 62.0%; Score 1992.5; DB 12; Length 623;  
 Best Local Similarity 61.4%; Pred. No. 1.7e-181;  
 Matches 378; Conservative 91; Mismatches 142; Indels 5; Gaps 3;

Qy 4 TTDYPLKNIHTPADIRALSKOQLQDLADVRGILTHYTSISGGHFAAGLGVETLVAAH 63  
 Db 5 STYFPLRSIQITDDIRTEBEADLTAVADELRAYLTESYKSGHFAAGLGVETLVAAH 64  
 Qy 64 YVNTPTDOLVMDVGHQAVPHKILNCRKEMPTIRTLGVSAPFPADESEYDAFGVGHSS 123  
 Db 65 YVQTPVDOLVMDVGHQAVPHKILNCRKEMPTIRTLGVSAPFPADESEYDAFGVGHSS 124  
 Qy 124 TSISALGMAIASQLGEGDKKVAITIGDGSITGGMAVEANNAAGDVANLVLINDMSISP 181  
 Db 125 TSISALGMAIAAQRGDDRRKVAIVIGDAMTAGMYEALNNAAGDPPNLLVILINDR 184  
 Qy 182 MSISPPVGMNNYLTLYVSSKFSVRESKKAALAKMPSVWEILARKTEEHVKGIVTGLFEELG 239  
 Db 185 MSISPPVGMNNYLTLYVSSKFSVRESKKAALAKMPSVWEILARKTEEHVKGIVTGLFEELG 244  
 Qy 240 TLFEEIGFNYEGPIDGHDVEMVSTLENLKDILTGPFVLHVTKKGQVAPAEKDPVAVGPAFDP 299  
 Db 245 TLFEEIGFNYEGPIDGHDVEMVSTLENLKDILTGPFVLHVTKKGQVAPAEKDPVAVGPAFDP 304  
 Qy 300 VPAFDPDKPLPRAABSPHTTYEVGKMLCDMAADDERLGLTPMREGSGLVESQTFPRRYDV 365  
 Db 305 VGPDPDSKGLVAAAG -AKPRTYDVSDVCDMAADDERLGLTPMREGSGLVESQTFPRRYDV 366  
 Qy 360 PNRYPVVALAEGQAVTLAAGQAQOGAKPVVAISTFLQSGYDQLHDVALQNLMDLFDLDRAGLVP 419  
 Db 364 PNRYPVVALAEGQAVTLAAGQAQOGAKPVVAISTFLQSGYDQLHDVALQNLMDLFDLDRAGLVP 423  
 Qy 420 RAGLVGDDGPTTHAGADPYSIMRCIPNMLMAADENECQMLTTGQHH -GPAASVRYPPKGPAAI 479  
 Db 424 RAGLVGDDGPTTHAGADPYSIMRCIPNMLMAADENECQMLTTGQHH -GPAASVRYPPKGPAAI 483  
 Qy 480 GPPAALDPLTALIEIGKAEVRHSGRIALANGSWYTPAVBAGKOLGATVVMREVPKFPDAILB 539  
 Db 484 GPPAALDPLTALIEIGKAEVRHSGRIALANGSWYTPAVBAGKOLGATVVMREVPKFPDAILB 543  
 Qy 540 QALVLEIARTHDVFTVEENVVIAAGAGSAINTFLOAQKVLMPVCNIGLPDRFVBOGSREELSLVG 599  
 Db 544 XAMLEIARTHDVFTVEENVVIAAGAGSAINTFLOAQKVLMPVCNIGLPDRFVBOGSREELSLVG 603  
 Qy 600 LLSLVGIDSKGLIATI 615

Db 604 LIAEAGIDGAGIRAAV 619

```

RESULT 13
US-10-369-493-16281
/ Sequence 16281, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ PRIOR FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 16281
/ LENGTH: 623
/ TYPE: PRT
/ ORGANISM: Xanthomonas campestris
US-10-369-493-16281

```

Query Match 62.0%; Score 1992.5; DB 12; Length 623;  
 Best Local Similarity 61.4%; Pred. No. 1.7e-181;  
 Matches 378; Conservative 91; Mismatches 142; Indels 5; Gaps 3;

QY 4 TTDYPLKXHTPADIRALSKDQQLADEVRYGTHVTSISGGHFAAGLGVETVALH 63  
 Db 5 STRYPLRSITQTPDLRFEENDLRAVDELRAVLESVSGSGHFAAGLGVETVALH 64  
 QY 64 YVFNTPVDQVWDVGHQAVPHKILTRERKMPRTITLGGVSAFPARDESEYDAFGVSHS 123  
 Db 65 YVGFPEVDQVWDVGHQAVPHKILTRERDQITVAKQGVAPFPREBSVYDTGVGSHS 124  
 QY 124 TSISALAGALIASQLRGEDKKNVAILIGGSLTGAAVYANAHAG--DYNAALVILINDND 181  
 Db 125 TSISALAGALIAQRGDRKVAVIGGAMTAGVYALAHAGGMDPEPMLVILINDNR 184  
 QY 182 MSISPPVGMNNYLTFKVLSKFSVRSBSKKALA--TPPSVWEIARKTEEHVAKMIVPG 239  
 Db 185 MSISEAVGLTMDLGRASGSRITLNAIRGGKKILCDKKNPTARFVRWEHWMGMFVPS 244  
 QY 240 TLFEBELGNYFGPIGDHVENMLVSTLENLKDLTGVEFLHVTYTKGKGTAPEKOPLAYHG 299  
 Db 245 TLFEBMGFTYTPIDGHDLPSLVGALKTLKTLGQQLHVTYTKGKGTELEAGDQIGYHA 304  
 QY 300 VFAPPTKDFLKPAAKPSHPHTTEVEGSMCDMAAODERLIGTFPAMEBSGGLVEFSQKF 359  
 Db 305 VGFPPDSKGLVAKAG-AKGPYTTDVFSWVCMMAADPKMLVITPAMEBSGGLVRFSEY 363  
 QY 360 PNEYFDVAIAEQHAVTLLAAGACOGAKPVVAIYSTFLOGYDQILHVDALQNLMDLFAID 419  
 Db 364 PORYFDVAIAEQHAVTLLAAGAAVTAQAKPVVAIYSTFLOGYDQILHVDAAVKDLVLFAD 423  
 QY 420 RAGLVGPDGPTAGAFDVSYRCIPNMLIMAPADNEBCROMLTTFQHHGASVRYPRGK 479  
 Db 424 RGVGVGPDGATAGALDLSFLACVHPHVMYMAADABRCOMLTTLGAREGHAARYPRGT 483  
 QY 480 GPGALIDPTLALISGKAEVRHSGRIAILAMGSWNTPAVEAGKQLGATVNNKPFVKFD 539  
 Db 484 GGTMLDASLTTLPIGKQOLRHSGARIALLGATVDAABAAGVRELGITLVNNMFVYKLD 543  
 QY 540 QALVLEIARTHDVFTVEENVTAGAGSAINFTLDAQVYLMPCVNCIGLPDFVYGGSSRE 599  
 Db 544 KMLEFLAKCHAPFVSIEDNVVAGAGSVSELLNAESVIMPMHLGLPDFQHAASRED 603  
 QY 600 LSLVGLDSKGIILATI 615

Db 604 LIAEAGIDGAGIRAAV 619

```

RESULT 14
US-10-369-493-15540
/ Sequence 15540, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ PRIOR FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 15540
/ LENGTH: 653
/ TYPE: PRT
/ ORGANISM: Xanthomonas campestris
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(653)
/ OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-15540

```

Query Match 62.0%; Score 1992.5; DB 12; Length 653;  
 Best Local Similarity 61.4%; Pred. No. 1.8e-181;  
 Matches 378; Conservative 91; Mismatches 142; Indels 5; Gaps 3;

QY 4 TTDYPLKXHTPADIRALSKDQQLADEVRYGTHVTSISGGHFAAGLGVETVALH 63  
 Db 27 STRYPLRSITQTPDLRFEENDLRAVDELRAVLESVSGSGHFAAGLGVETVALH 96  
 QY 64 YVFNTPVDQVWDVGHQAVPHKILTRERKMPRTITLGGVSAFPARDESEYDAFGVSHS 123  
 Db 87 YVGFPEVDQVWDVGHQAVPHKILTRERDQITVAKQGVAPFPREBSVYDTGVGSHS 146  
 QY 124 TSISALAGALIASQLRGEDKKNVAILIGGSLTGAAVYANAHAG--DYNAALVILINDND 181  
 Db 147 TSISALAGALIAQRGDRKVAVIGGAMTAGVYALAHAGGMDPEPMLVILINDNR 206  
 QY 182 MSISPPVGMNNYLTFKVLSKFSVRSBSKKALA--TPPSVWEIARKTEEHVAKMIVPG 239  
 Db 207 MSISEAVGLTMDLGRASGSRITLNAIRGGKKILCDKKNPTARFVRWEHWMGMFVPS 266  
 QY 240 TLFEBELGNYFGPIGDHVENMLVSTLENLKDLTGVEFLHVTYTKGKGTAPEKOPLAYHG 299  
 Db 267 TLFEBMGFTYTPIDGHDLPSLVGALKTLKTLGQQLHVTYTKGKGTELEAGDQIGYHA 326  
 QY 300 VFAPPTKDFLKPAAKPSHPHTTEVEGSMCDMAAODERLIGTFPAMEBSGGLVEFSQKF 359  
 Db 327 VGFPPDSKGLVAKAG-AKGPYTTDVFSWVCMMAADPKMLVITPAMEBSGGLVRFSEY 385  
 QY 360 PNEYFDVAIAEQHAVTLLAAGACOGAKPVVAIYSTFLOGYDQILHVDALQNLMDLFAID 419  
 Db 386 PORYFDVAIAEQHAVTLLAAGAAVTAQAKPVVAIYSTFLOGYDQILHVDAAVKDLVLFAD 445  
 QY 420 RAGLVGPDGPTAGAFDVSYRCIPNMLIMAPADNEBCROMLTTFQHHGASVRYPRGK 479  
 Db 446 RGVGVGPDGATAGALDLSFLACVHPHVMYMAADABRCOMLTTLGAREGHAARYPRGT 505  
 QY 480 GPGALIDPTLALISGKAEVRHSGRIAILAMGSWNTPAVEAGKQLGATVNNKPFVKFD 539  
 Db 506 GGTMLDASLTTLPIGKQOLRHSGARIALLGATVDAABAAGVRELGITLVNNMFVYKLD 565  
 QY 540 QALVLEIARTHDVFTVEENVTAGAGSAINFTLDAQVYLMPCVNCIGLPDFVYGGSSRE 599

Db 566 KAMELELAKCHEAFISIDNVAGAGSGVSELLNABSVLPMHLGLPDSFOHARESD 625  
Qy 600 LLSLVGDSKGLIATI 615  
Db 626 LLAEXGIDQAGIRAAV 641

Db 542 LDADLVRLAETHDAIVTEBGCWAGGSACTVEALLASGVTRPVQLGLPDRFIDHSDP 601  
Qy 598 BELLSLVGDSKGLIATI 615  
Db 602 AKLLAACGIDAVGITKSI 619

Search completed: January 29, 2004, 16:21:14  
Job time : 44.325 secs

RESULT 15  
US-10-369-493-7428  
; Sequence 7428, Application US/10369493  
; Publication No. US2003023675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xiaofeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 7428  
; LENGTH: 619  
; TYPE: PRN  
; ORGANISM: Burkholderia cepacia  
US-10-369-493-7428

Query Match 61.9%; Score 1990.5; DB 12; Length 619;  
Best Local Similarity 61.7%; Pred. No. 2.6e-181;  
Matches 381; Conservative 89; Mismatches 141; Indels 7; Gaps 3;

Qy 4 TTDVPLKNIHTPADIRALSKDQLOQLADEVRGYLTHVTSISGHPAAGTVELVLAH 63  
Db 3 TTMVOLLKTIIDPPALRLDRQLPLADELRAPLDSVSTGCHLSNLTGVELTIALH 62  
Qy 64 YVFNTVPDQVMDVGHQAYPHKILTKRERMPITRTLGVSAPPARDBSEYDAFGVHSS 123  
Db 63 YVFDTPHDRIVMDVGHQYTPHKILTKGRDQMTLRLQLGISGFPKRDESEYDTPTAASS 122  
Qy 124 TSSAALGMALISOLRGEDKKONATIGSGITGKAYEAMNHAG-DVNAANLVLINDNM 182  
Db 123 TSSALGMANASKLGKGNRMGIAVIGGANTAGCAFPAMNNAAGVEDVFLVILINDNM 182  
Qy 183 SISPPVGMANNLYTVLSSKPYSSVREBSKALAMPSEWELARKTEBHVKGMIVPGLTF 242  
Db 183 SISPPVGMANNLYTVLSSKPYSSVREBSKALAMPSEWELARKTEBHVKGMIVPGLTF 242  
Qy 243 BELGNNYRGPIRGHVEMVSTLENLKDLPVPLHYVTKKGYAPAEKDLAYHGVPA 302  
Db 243 BELGNNYRGPIRGHVEMVSTLENLKDLPVPLHYVTKKGYAPAEKDLAYHGVPA 302  
Qy 243 BELGNNYRGPIRGHVEMVSTLENLKDLPVPLHYVTKKGYAPAEKDLAYHGVPA 302  
Db 243 BELGNNYRGPIRGHVEMVSTLENLKDLPVPLHYVTKKGYAPAEKDLAYHGVPA 302  
Qy 303 PDPTDPLPKAAPSPHYTEVFEGRMLCDMAODERLLGITPAVBEGSGVFEFSQKFPNR 362  
Db 303 PDPTDPLPKAAPSPHYTEVFEGRMLCDMAODERLLGITPAVBEGSGVFEFSQKFPNR 362  
Qy 363 YFDVVAIAEGHATVTLAAGQACGAKPVAAIYSTFLORGIDQIHDVALQNLDMFLALDRAG 422  
Db 363 YFDVVAIAEGHATVTLAAGQACGAKPVAAIYSTFLORGIDQIHDVALQNLDMFLALDRAG 422  
Qy 423 LVGPDGPTAGAFDYSYMRCTPMLIMAPADENECQMLTTFQHHGPAVSAYPRGKPG 482  
Db 423 LVGPDGPTAGAFDYSYMRCTPMLIMAPADENECQMLTTFQHHGPAVSAYPRGKPG 482  
Qy 483 AALDPTLALBETGKAEVRRH-----GSRIALIANGSMTPAVERAKQIAGATVVMRFYKP 537  
Db 483 AALDPTLALBETGKAEVRRH-----GSRIALIANGSMTPAVERAKQIAGATVVMRFYKP 537  
Qy 538 PQGALVLELARTHDVFTVVEENVVAGGASAINTEFLQAKVLMPCNIGLPDRFVEQGSR 597  
Db 538 PQGALVLELARTHDVFTVVEENVVAGGASAINTEFLQAKVLMPCNIGLPDRFVEQGSR 597

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:37:58 ; Search time 39 Seconds  
(without alignments)

2328.695 Million cell updates/sec

Title: US-09-941-947a-2

Perfect score: 2284

Sequence: 1 DVVTWPHLTADIRFCWFF.....FNMELSRIGQKGRKVEVSH 437

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCR\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCRUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2284	100.0	437	10	US-09-934-901-16
2	2284	100.0	437	10	US-09-934-868-6
3	2284	100.0	437	11	US-09-941-947a-2
4	2284	100.0	437	12	US-10-320-874-16
5	2284	100.0	437	16	US-10-320-924-16
6	659	28.9	184	12	US-10-369-493-14348
7	306.5	13.4	350	12	US-10-369-493-8108
8	294	12.9	342	15	US-10-156-761-13610
9	284.5	12.5	341	15	US-10-156-761-14652
10	269.5	11.8	341	15	US-10-156-761-10359
11	261	11.4	345	12	US-10-369-493-19761
12	253	11.1	319	12	US-10-369-493-2905
13	250.5	11.0	357	12	US-10-369-493-9669
14	246.5	10.8	987	12	US-10-369-493-21998
15	240.5	10.5	320	9	US-09-815-242-13785

16	239.5	10.5	318	12	US-10-369-493-9652	Sequence 9652, Ap
17	238.5	10.4	320	9	US-09-815-242-10422	Sequence 10422, A
18	238.5	10.4	320	12	US-10-369-493-23638	Sequence 23638, A
19	229.5	10.0	295	12	US-10-369-493-8855	Sequence 8855, Ap
20	227.5	10.0	449	12	US-10-369-493-8876	Sequence 8876, Ap
21	226.5	9.9	320	9	US-09-815-242-11721	Sequence 11721, A
22	224.5	9.8	334	12	US-10-369-493-515	Sequence 515, Ap
23	223.5	9.8	319	12	US-10-369-493-17399	Sequence 17399, A
24	222.5	9.7	322	9	US-09-815-242-5425	Sequence 5425, Ap
25	222.5	9.7	322	9	US-09-815-242-12619	Sequence 12619, A
26	222.5	9.7	322	9	US-09-815-242-12907	Sequence 12907, A
27	222.5	9.7	360	12	US-10-369-493-19198	Sequence 19198, A
28	222	9.7	322	12	US-10-369-493-10132	Sequence 10132, A
29	222	9.7	340	12	US-10-369-493-18431	Sequence 18431, A
30	221	9.7	335	9	US-09-815-242-13346	Sequence 13346, A
31	218.5	9.6	320	9	US-09-815-242-10649	Sequence 10649, A
32	215	9.4	321	12	US-10-369-493-21210	Sequence 21210, A
33	213.5	9.3	321	12	US-10-369-493-128	Sequence 128, Ap
34	213	9.3	321	12	US-10-369-493-479	Sequence 479, Ap
35	207	9.1	959	12	US-10-369-493-1903	Sequence 1903, Ap
36	202.5	8.9	321	9	US-09-815-242-11162	Sequence 11162, A
37	201.5	8.8	361	12	US-10-369-493-2625	Sequence 2625, Ap
38	201	8.8	343	9	US-09-887-054-2	Sequence 2, Appl
39	201	8.8	343	15	US-10-098-626-2	Sequence 2, Appl
40	201	8.8	357	10	US-09-738-626-4883	Sequence 4883, Ap
41	200	8.8	776	15	US-10-305-342-5	Sequence 5, Appl
42	198	8.7	419	12	US-10-369-493-2919	Sequence 2919, Ap
43	196	8.6	321	12	US-10-369-493-10434	Sequence 10434, A
44	196	8.6	828	12	US-10-369-493-6730	Sequence 6730, Ap
45	193	8.5	942	12	US-10-369-493-2497	Sequence 2497, Ap

## ALIGNMENTS

US-09-934-901-16	Sequence 16, Application US/09934901
Patent No. US20020110885A1	
GENERAL INFORMATION	
APPLICANT: scifess, Matthews	
APPLICANT: Odum, J. Mattheos	
APPLICANT: No. US20020110885A10m, Kelley C.	
APPLICANT: Ye, Rick	
TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN	
FILE REFERENCE: C14619 US NA	
CURRENT APPLICATION NUMBER: US/09/934,901	
PRIOR APPLICATION NUMBER: 60/223,906	
PRIOR FILING DATE: September 1, 2000	
NUMBER OF SEQ ID NOS: 20	
SOFTWARE: Microsoft Office 97	
SEQ ID NO 16	
LENGTH: 437	
TYPE: PRT	
ORGANISM: METHYLOMONAS SP.	
US-09-934-901-16	
Query Match	100.0%; Score 2284; DB 10; Length 437;
Best Local Similarity	100.0%; Pred. No. 1.4e-223;
Matches	437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 DVVTWPHLTADIRFCWFFLNRRFYTLNKKPKKVAALITGGLAPCLNSAIGSLIRYTE 60
DB	1 DVVTWPHLTADIRFCWFFLNRRFYTLNKKPKKVAALITGGLAPCLNSAIGSLIRYTE 60
QY	61 IDPFIETICRGYKGLLSDSYFVTAVERKKAGVQRFQGSYIGNSRVLTAVKQCVKR 120
DB	61 IDPFIETICRGYKGLLSDSYFVTAVERKKAGVQRFQGSYIGNSRVLTAVKQCVKR 120
QY	121 GLVTEGEDPQKVAADQVLKQGVDLHTIGGDDNTYAAADLAATLARNRYGLTYIGPKTY 180
DB	121 GLVTEGEDPQKVAADQVLKQGVDLHTIGGDDNTYAAADLAATLARNRYGLTYIGPKTY 180

Query 181 DNDVFPKOSIGANTAAOGARFENNVVAENNNANPRLIVHEVWGNCGULTAATAOEYR 240  
Db 181 DNDVFPKOSIGANTAAOGARFENNVVAENNNANPRLIVHEVWGNCGULTAATAOEYR 240  
Query 241 KLDRAEWLPELGLTRSYEVHAFVPEMAIDLEAEKRLREVMKYDCVNIIVSAGAV 300  
Db 241 KLDRAEWLPELGLTRSYEVHAFVPEMAIDLEAEKRLREVMKYDCVNIIVSAGAV 300  
Query 301 EAIYAEMOAKGOEYPRDAFGHILKIDAVNPGKMGEOPOAOMIGAKTLVOKSGYFARASAS 360  
Db 301 EAIYAEMOAKGOEYPRDAFGHILKIDAVNPGKMGEOPOAOMIGAKTLVOKSGYFARASAS 360  
Query 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420  
Db 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420  
Query 421 MLESEIGOPKGGKXEVSH 437  
Db 421 MLESEIGOPKGGKXEVSH 437

RESULT 2  
US-09-934-868-6  
Sequence 6, Application US/09934868  
Patent No. US20020137190A1  
GENERAL INFORMATION:  
APPLICANT: Koffas, Mattheos  
APPLICANT: Odom, James M  
APPLICANT: Schenzle, Andreas J  
TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN  
FILE REFERENCE: C11596 US NA  
CURRENT APPLICATION NUMBER: US/09/934,868  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/229,858  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 6  
LENGTH: 437  
TYPE: PRT  
ORGANISM: METHYLOMONAS SP.  
US-09-934-868-6

Query Match 100.0%; Score 2284; DB 10; Length 437;  
Best Local Similarity 100.0%; Pred. No. 1.4e-223;  
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 DVTWPHYLTADIRFCMFPLNFYTLNKPCKKVALITAGGLAPCLNSAIGSLIRYTE 60  
Db 1 DVTWPHYLTADIRFCMFPLNFYTLNKPCKKVALITAGGLAPCLNSAIGSLIRYTE 60  
Query 61 IDPSIIEICRGYKGLLDSDYPVTAEVKKAQVLOREFGSVI GNSRVLTNVKDCVKR 120  
Db 61 IDPSIIEICRGYKGLLDSDYPVTAEVKKAQVLOREFGSVI GNSRVLTNVKDCVKR 120  
Query 121 GLVEGEDPOKVAADQVJKGVLDLHTTGGDDNTAAADLAFLAENNYGLTVIGLPKTV 180  
Db 121 GLVEGEDPOKVAADQVJKGVLDLHTTGGDDNTAAADLAFLAENNYGLTVIGLPKTV 180  
Query 181 DNDVFPKOSIGANTAAOGARFENNVVAENNNANPRLIVHEVWGNCGULTAATAOEYR 240  
Db 181 DNDVFPKOSIGANTAAOGARFENNVVAENNNANPRLIVHEVWGNCGULTAATAOEYR 240  
Query 241 KLDRAEWLPELGLTRSYEVHAFVPEMAIDLEAEKRLREVMKYDCVNIIVSAGAV 300  
Db 241 KLDRAEWLPELGLTRSYEVHAFVPEMAIDLEAEKRLREVMKYDCVNIIVSAGAV 300  
Query 301 EAIYAEMOAKGOEYPRDAFGHILKIDAVNPGKMGEOPOAOMIGAKTLVOKSGYFARASAS 360  
Db 301 EAIYAEMOAKGOEYPRDAFGHILKIDAVNPGKMGEOPOAOMIGAKTLVOKSGYFARASAS 360  
Query 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420  
Db 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420

Query 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420  
Db 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420  
Query 421 MLESEIGOPKGGKXEVSH 437  
Db 421 MLESEIGOPKGGKXEVSH 437

RESULT 3  
US-09-941-947a-2  
Sequence 2, Application US/09941947a  
Publication No. US20030003528A1  
GENERAL INFORMATION:  
APPLICANT: Bzostowicz, Patricia C.  
APPLICANT: Cheng, Qiong  
APPLICANT: DiCosimo, Deana J.  
APPLICANT: Koffas, Mattheos  
APPLICANT: Miller, Edward S. Jr.  
APPLICANT: Odom, J. Martin  
APPLICANT: Picataggio, Steve  
APPLICANT: Rouviere, Pierre E.  
TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE  
FILE REFERENCE: C11903 US NA  
CURRENT APPLICATION NUMBER: US/09/941,947A  
CURRENT FILING DATE: 2001-09-01  
PRIOR APPLICATION NUMBER: 60/229,907  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229,858  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 60  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 2  
LENGTH: 437  
TYPE: PRT  
ORGANISM: Methylomonas 16a  
US-09-941-947a-2

Query Match 100.0%; Score 2284; DB 11; Length 437;  
Best Local Similarity 100.0%; Pred. No. 1.4e-223;  
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 DVTWPHYLTADIRFCMFPLNFYTLNKPCKKVALITAGGLAPCLNSAIGSLIRYTE 60  
Db 1 DVTWPHYLTADIRFCMFPLNFYTLNKPCKKVALITAGGLAPCLNSAIGSLIRYTE 60  
Query 61 IDPSIIEICRGYKGLLDSDYPVTAEVKKAQVLOREFGSVI GNSRVLTNVKDCVKR 120  
Db 61 IDPSIIEICRGYKGLLDSDYPVTAEVKKAQVLOREFGSVI GNSRVLTNVKDCVKR 120  
Query 121 GLVEGEDPOKVAADQVJKGVLDLHTTGGDDNTAAADLAFLAENNYGLTVIGLPKTV 180  
Db 121 GLVEGEDPOKVAADQVJKGVLDLHTTGGDDNTAAADLAFLAENNYGLTVIGLPKTV 180  
Query 181 DNDVFPKOSIGANTAAOGARFENNVVAENNNANPRLIVHEVWGNCGULTAATAOEYR 240  
Db 181 DNDVFPKOSIGANTAAOGARFENNVVAENNNANPRLIVHEVWGNCGULTAATAOEYR 240  
Query 241 KLDRAEWLPELGLTRSYEVHAFVPEMAIDLEAEKRLREVMKYDCVNIIVSAGAV 300  
Db 241 KLDRAEWLPELGLTRSYEVHAFVPEMAIDLEAEKRLREVMKYDCVNIIVSAGAV 300  
Query 301 EAIYAEMOAKGOEYPRDAFGHILKIDAVNPGKMGEOPOAOMIGAKTLVOKSGYFARASAS 360  
Db 301 EAIYAEMOAKGOEYPRDAFGHILKIDAVNPGKMGEOPOAOMIGAKTLVOKSGYFARASAS 360  
Query 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420  
Db 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420  
Query 421 MLESEIGOPKGGKXEVSH 437  
Db 421 MLESEIGOPKGGKXEVSH 437

RESULT 4  
 US-10-320-874-16  
 ; Sequence 16, Application US/10320874  
 ; Publication No. US20030138909A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kofess, Mattheos  
 ; APPLICANT: Odou, J. Martin  
 ; APPLICANT: No. US20030138909A1ton, Kelley C.  
 ; APPLICANT: Ye, Rick  
 ; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN  
 ; FILE REFERENCE: CL1619 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/320,874  
 ; CURRENT FILING DATE: 2002-12-17  
 ; PRIOR APPLICATION NUMBER: US/09/934,901  
 ; PRIOR FILING DATE: 2001-08-22  
 ; PRIOR APPLICATION NUMBER: 60/229,906  
 ; PRIOR FILING DATE: September 1, 2000  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 16  
 ; LENGTH: 437  
 ; TYPE: PRT  
 ; ORGANISM: METHYLOMONAS SP.  
 US-10-320-874-16

Query Match 100.0%; Score 2284; DB 12; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-223;  
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DVTWPHLTADIRCHWFLENFYTLMNPKKVAIITAGGLAPCLNSAIGSLIERYTE	60
DB	1	DVTWPHLTADIRCHWFLENFYTLMNPKKVAIITAGGLAPCLNSAIGSLIERYTE	60
QY	61	IDPSIEIICRGYKGLLDGSDYPVTAVERKAGVLRFGGSVIGNSRVKLTNVKDCVCR	120
DB	61	IDPSIEIICRGYKGLLDGSDYPVTAVERKAGVLRFGGSVIGNSRVKLTNVKDCVCR	120
QY	121	GLVKEGDPORVADOLVKDGVDLHTIGGDDITAAADLAFLARNNGYGLTVGLPKTV	180
DB	121	GLVKEGDPORVADOLVKDGVDLHTIGGDDITAAADLAFLARNNGYGLTVGLPKTV	180
QY	181	DNDVFPKOSIGAMTAAGQARYFMNVVAENNANPRLIYHEVWGRNCGMLTAATAOYR	240
DB	181	DNDVFPKOSIGAMTAAGQARYFMNVVAENNANPRLIYHEVWGRNCGMLTAATAOYR	240
QY	241	KLDRAEWLPELGLTRSEYVHAVFPEMAIDLEBAKRLREVMDKVDVNIIVSEAGV	300
DB	241	KLDRAEWLPELGLTRSEYVHAVFPEMAIDLEBAKRLREVMDKVDVNIIVSEAGV	300
QY	301	EAIYAEWQAKQOEVRDPAFGHIKLDVAVPGKMFGEOPAQMTGAETLVQSGYFARASAS	360
DB	301	EAIYAEWQAKQOEVRDPAFGHIKLDVAVPGKMFGEOPAQMTGAETLVQSGYFARASAS	360
QY	361	NVDWRLIKSCADLAABCAFRRESGVIGHEDNGNVLRAIEFPRIKGGKPNIDTWMNS	420
DB	361	NVDWRLIKSCADLAABCAFRRESGVIGHEDNGNVLRAIEFPRIKGGKPNIDTWMNS	420
QY	421	MLSEIGOPKGGKVEVSH 437	
DB	421	MLSEIGOPKGGKVEVSH 437	

RESULT 5  
 US-10-320-924-16  
 ; Sequence 16, Application US/10320924  
 ; Publication No. US20030129721A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kofess, Mattheos  
 ; APPLICANT: Odou, J. Martin  
 ; APPLICANT: No. US20030129721A1ton, Kelley C.  
 ; APPLICANT: Ye, Rick  
 ; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN

FILE REFERENCE: CL1619 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/320,924  
 ; CURRENT FILING DATE: 2002-12-17  
 ; PRIOR APPLICATION NUMBER: US/09/934,901  
 ; PRIOR FILING DATE: 2001-08-22  
 ; PRIOR APPLICATION NUMBER: 60/229,906  
 ; PRIOR FILING DATE: September 1, 2000  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 16  
 ; LENGTH: 437  
 ; TYPE: PRT  
 ; ORGANISM: METHYLOMONAS SP.  
 US-10-320-924-16

Query Match 100.0%; Score 2284; DB 16; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-223;  
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DVTWPHLTADIRCHWFLENFYTLMNPKKVAIITAGGLAPCLNSAIGSLIERYTE	60
DB	1	DVTWPHLTADIRCHWFLENFYTLMNPKKVAIITAGGLAPCLNSAIGSLIERYTE	60
QY	61	IDPSIEIICRGYKGLLDGSDYPVTAVERKAGVLRFGGSVIGNSRVKLTNVKDCVCR	120
DB	61	IDPSIEIICRGYKGLLDGSDYPVTAVERKAGVLRFGGSVIGNSRVKLTNVKDCVCR	120
QY	121	GLVKEGDPORVADOLVKDGVDLHTIGGDDITAAADLAFLARNNGYGLTVGLPKTV	180
DB	121	GLVKEGDPORVADOLVKDGVDLHTIGGDDITAAADLAFLARNNGYGLTVGLPKTV	180
QY	181	DNDVFPKOSIGAMTAAGQARYFMNVVAENNANPRLIYHEVWGRNCGMLTAATAOYR	240
DB	181	DNDVFPKOSIGAMTAAGQARYFMNVVAENNANPRLIYHEVWGRNCGMLTAATAOYR	240
QY	241	KLDRAEWLPELGLTRSEYVHAVFPEMAIDLEBAKRLREVMDKVDVNIIVSEAGV	300
DB	241	KLDRAEWLPELGLTRSEYVHAVFPEMAIDLEBAKRLREVMDKVDVNIIVSEAGV	300
QY	301	EAIYAEWQAKQOEVRDPAFGHIKLDVAVPGKMFGEOPAQMTGAETLVQSGYFARASAS	360
DB	301	EAIYAEWQAKQOEVRDPAFGHIKLDVAVPGKMFGEOPAQMTGAETLVQSGYFARASAS	360
QY	361	NVDWRLIKSCADLAABCAFRRESGVIGHEDNGNVLRAIEFPRIKGGKPNIDTWMNS	420
DB	361	NVDWRLIKSCADLAABCAFRRESGVIGHEDNGNVLRAIEFPRIKGGKPNIDTWMNS	420
QY	421	MLSEIGOPKGGKVEVSH 437	
DB	421	MLSEIGOPKGGKVEVSH 437	

RESULT 6  
 US-10-369-493-14348  
 ; Sequence 14348, Application US/10369493  
 ; Publication No. US2003023675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 14348  
 ; LENGTH: 184  
 ; TYPE: PRT

ORGANISM: Agrobacterium tumefaciens  
US-10-369-493-13348

Query Match 28.9%; Score 659; DB 12; Length 184;  
Best Local Similarity 67.9%; Pred. No. 1.1e-58;  
Matches 125; Conservative 27; Mismatches 32; Indels 0; Gaps 0;

QY 33 KXVALITAGGLAPCLNSAIGSLERTEIDPSIEIICYGKYGKLLGDSVPVTAVERK 92  
DB 1 QXVAMLTGGLAPCLNSAIGSLERTEIDPSIEIICYGKYGKLLGDSVPVTAVERK 60  
QY 93 AGVLOPFGSVIGNSRYKLTNVKDCVKGGLVKEGEDPQKVAADQVLDLHTIGDD 152  
DB 61 AHIHRYGSGPIGNSRYKLTNVKDCVKGGLVKEGEDPQKVAADQVLDLHTIGDD 120  
QY 153 TMTAAADLAFLAANNYGLTVIGLPTVNDVPPIKOSLGAMTAAGARVPMVAENN 212  
DB 121 TMTAAADLAFLAANNYGLTVIGLPTVNDVPPIKOSLGAMTAAGARVPMVAENN 180  
QY 213 ANPR 216  
DB 181 AAKR 184

## RESULT 7

US-10-369-493-8108  
Sequence 8108, Application US/10369493  
Publication No. US2003023675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
PLANTS WITH IMPROVED PROPERTIES  
FILE REFERENCE: 38-10(52052)B  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US/10/369,493  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 60/360,039  
SEQ ID NO 8108  
LENGTH: 350  
TYPE: PRT  
ORGANISM: Thermobifida fusca  
US-10-369-493-8108

Query Match 13.4%; Score 306.5; DB 12; Length 350;  
Best Local Similarity 26.5%; Pred. No. 2.2e-22;  
Matches 104; Conservative 63; Mismatches 145; Indels 81; Gaps 14;

QY 23 FNYFTLANKPKKVALITAGGLAPCLNSAIGSLERTEIDPSIEIICYGKYGKLLGDS 82  
DB 4 FKEATQK---RVGLVLTGGDPCGLNAVIRAVVRKGLK-EYGYEFVGFDDMGKPLSGDT 58  
QY 83 YPVTAVERKXKAGVQRFPGSVIGNSRYKLTNVKDCVKGGLVKEGEDPQKVAADQVLDV 142  
DB 59 MFLIEAVR-GILPR-GGITLSSRTNLKIBGVYR-----VDNMAALAV 103  
QY 143 DILHTIGGDDTNTAAADLAFLAANNYGLTVIGLPTVNDVPPIKOSLGAMTAAGAR 202  
DB 104 DALVAIGGEDTLGVARQI-----HDGVAVVGVFKTIDNDLNAATYTFGFDVAVIAIE 157  
QY 203 YENNVAVENNANPEYLLVHEVMGRNGMLT-----ATAQGYKLLDRAEMLPELGITRE 257  
DB 158 AIDRLHTTAESEHRAVY-EEVGHAGKIALHAGMAGA----- 195  
QY 258 SEYVAHVFPPEMAIDLEABARLEEVNDKVCNIPVSEGAQEALVAENQAKGEVPRD 317  
DB 196 ---NVILIFRPRPDIDEVVAVYIESRFTNVAPIIVABAHK---EGQLTAGAERD 247  
QY 318 AFGHIKLDVAVNGMFGEGFQMGIMGENTVQKS---GYRARSASNVDMRLKSCADL 374

DB 248 SFGHVRVAGI-----GQRLAEIEARTKEARSVLGHVQCGTSPAFDRVLAIFLGLH 301  
QY 375 AVECAFRREGVIGHDEDNQNV--LRAISFPRI 405  
DB 302 ALTRV-----HDKDFKQVALLGTHIVRV 325

## RESULT 8

US-10-156-761-13610  
Sequence 13610, Application US/10156761  
Publication No. US20030119018A1  
GENERAL INFORMATION:  
APPLICANT: OMIYA, SATOSHI  
APPLICANT: IKEDA, HARUO  
APPLICANT: ISHIKAWA, JUN  
APPLICANT: HORIKAWA, HIROSHI  
APPLICANT: SHIBA, TADAYOSHI  
APPLICANT: SAKAKI, YOSHIYUKI  
APPLICANT: HATTORI, MASAHIRA  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-262  
CURRENT APPLICATION NUMBER: US/10/156,761  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: JP 2001-204089  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: JP 2001-272697  
PRIOR FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 15109  
SEQ ID NO 13610  
LENGTH: 342  
TYPE: PRT  
ORGANISM: Streptomyces avermitilis  
US-10-156-761-13610

Query Match 12.9%; Score 294; DB 15; Length 342;  
Best Local Similarity 26.5%; Pred. No. 4e-21;  
Matches 101; Conservative 66; Mismatches 140; Indels 74; Gaps 17;

QY 34 KVALITGGLAPCLNSAIGSLERTEIDPSIEIICYGKYGKLLGDSVPVTAVERKXKA 93  
DB 2 RVGLVLTGGDPCGLNAVIRAVVRKGVQ-EYGYDFVGFDDMGKPLSGDA--VALDPAVR 58  
QY 94 GVLQRFPGSVIGNSRYKLTNVKDCVKGGLVKEGEDPQKVAADQVLDLHTIGDDT 153  
DB 59 GILPR-GGITLSSRTNLKIBGVYR-----NLAKQVDAALIAIGSDT 105  
QY 154 MTAADLAFLAANNYGLTVIGLPTVNDVPPIKOSLGAMTAAGARV--MNVVAEN 211  
DB 106 LGVAARLT-----DEYGVPAVGVFKTIDNDLSATDYTFGFDVAVIAIEIDLHTTALS 160  
QY 212 NAMPKLTIVHEVMGRNGMLTAAATQOERYKLLDRAEMLPELGITRESYEVHAVFVPM 271  
DB 161 HFR---VLVEGVYGRHAGWIAHS-----GLAGCA---NVILIFEGQRF 197  
QY 272 DLEAARLEEVNDKVCNIPVSEGAQEALVAENQAKGEVPR---DAFHTLDAY 327  
DB 198 DVDQVCAVYVSRFASAPAVIVVABGA-----MFKQGMVLKQESLDSFHVRLSGV 249  
QY 328 NRGKMPGEPQAFQMGIAE-KTLVQKSGFAPASASNVDMRLKSCADLAVECAFRRBSGV 386  
DB 250 --GEMLAKEIKRGKGAERTTV--LGHVQCGTSPAFDRVLAIFLGLHIAEV----- 258  
QY 387 IGHDEDNQNV--LRAISFPRI 405  
DB 299 --RDGDFKQVALLGTHIVRV 317

RESULT 9  
US-10-156-761-14652  
Sequence 14652, Application US/10156761  
Publication No. US20030119018A1  
GENERAL INFORMATION:



```

; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14652
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14652

```

```

Query Match      12.5%; Score 284.5; DB 15; Length 341;
Best Local Similarity 27.5%; Pred. No. 3.7e-20;
Matches 97; Conservative 62; Mismatches 143; Indels 51; Gaps 11;

```

```

QY 34 KVAITTAGLAPCLNSAIGSLIERYTEIDPSIEIICRYGGYKGLLDGSPVTAEVKKA 93
DB 2 RIGVITAGGDCPGIAAVRSVVR-AVVDHGDVIGFRDGGKGLLECDY--LKDILAVS 58
QY 94 GVLORFGSGVIGNSRVKLTNVKDCVKGGLVEGEDPQKVAADQVKDGVILHTIGGDDT 153
DB 59 GILAR-GGTLIGSSSLQDRLREACEN-----AQDMAHREIDALIPIGSGGT 105
QY 154 NTAADLAFLARNNGITVIGLPTVNDVFPKOSIGAMTAABQARFMVVAENNA 213
DB 106 LKAARLL-----SDAGLPVGVPEKTIQNDIAVDVTFGPTAVGAVATEADRLKTTAES 159
QY 214 NPMILIVHVMGRNCGMTLTAATQERYKLDRAEWLPELGLTRSEYVHAFVPEMAIDL 273
DB 160 HORVWVW-EVMGRHAGWIAL-----ESGMAAGA--HGICLERPPDP 198
QY 274 EBAKRLAEVMDK-----VDCVNI FVSGAGVEALVAEMQKQGEVPRDAFGIKLDAVN 328
DB 199 AELVWVVERFARFGKPAVLC-----VAEGA--HPVGSMDYSHGAI--DQGHRRFOGI- 249
QY 329 PKWFGGEPQOMIGAE--TLVQKSGYFARASAVNDMRLIKSCADLAVECAFRRREG 385
DB 250 -GTALAAELERLIGKEARPV--LGHVORGSTPTAVDVLATRFGMHVAEVAHRRGFG 304

```

RESULT 10

```

US-10-156-761-10359
; Sequence 10359, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10359

```

```

; LENGTH: 341
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10359

```

```

Query Match      11.8%; Score 269.5; DB 15; Length 341;
Best Local Similarity 26.8%; Pred. No. 1.2e-18;
Matches 96; Conservative 59; Mismatches 142; Indels 61; Gaps 15;

```

```

QY 34 KVAITTAGLAPCLNSAIGSLIERYTEIDPSIEIICRYGGYKGLLDGSPVTAEVKKA 93
DB 2 RIGVITAGGDCPGIAAVRSVVR-AVVDHGDVIGFRDGGKGLLECDY--LKDILAVS 58
QY 94 GVLORFGSGVIGNSRVKLTNVKDCVKGGLVEGEDPQKVAADQVKDGVILHTIGGDDT 153
DB 59 GILAR-GGTLIGSSSLQDRLREACEN-----AQDMAHREIDALIPIGSGGT 105
QY 154 NTAADLAFLARNNGITVIGLPTVNDVFPKOSIGAMTAABQARFMVVAENNA 213
DB 106 LKAARLL-----SDAGLPVGVPEKTIQNDIAVDVTFGPTAVGAVATEADRLKTTAES 159
QY 214 NPMILIVHVMGRNCGMTLTAATQERYKLDRAEWLPELGLTRSEYVHAFVPEMAIDL 273
DB 160 HORVWVW-EVMGRHAGWIAL-----ESGMAAGA--HGICLERPPDP 198
QY 274 EBAKRLAEVMDK-----VDCVNI FVSGAGVEALVAEMQKQGEVPRDAFGIKLDAVN 328
DB 199 AELVWVVERFARFGKPAVLC-----VAEGA--HPVGSMDYSHGAI--DQGHRRFOGI- 249
QY 329 PKWFGGEPQOMIGAE--TLVQKSGYFARASAVNDMRLIKSCADLAVECAFRRREG 385
DB 250 -GTALAAELERLIGKEARPV--LGHVORGSTPTAVDVLATRFGMHVAEVAHRRGFG 304

```

RESULT 11

```

US-10-369-493-19261
; Sequence 19261, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19261
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-10-369-493-19261

```

```

Query Match      11.4%; Score 261; DB 12; Length 345;
Best Local Similarity 26.0%; Pred. No. 9.3e-18;
Matches 96; Conservative 66; Mismatches 141; Indels 66; Gaps 16;

```

```

QY 34 KVAITTAGLAPCLNSAIGSLIERYTEIDPSIEIICRYGGYKGLLDGSPVTAEVKKA 93
DB 2 RIGVITAGGDCPGIAAVRSVVR-AVVDHGDVIGFRDGGKGLLECDY--LKDILAVS 58
QY 94 GVLORFGSGVIGNSRVKLTNVKDCVKGGLVEGEDPQKVAADQVKDGVILHTIGGDDT 153
DB 59 GILAR-GGTLIGSSSLQDRLREACEN-----AQDMAHREIDALIPIGSGGT 105
QY 154 NTAADLAFLARNNGITVIGLPTVNDVFPKOSIGAMTAABQARFMVVAENNA 213
DB 106 LKAARLL-----SDAGLPVGVPEKTIQNDIAVDVTFGPTAVGAVATEADRLKTTAES 159
QY 214 NPMILIVHVMGRNCGMTLTAATQERYKLDRAEWLPELGLTRSEYVHAFVPEMAIDL 273
DB 160 HORVWVW-EVMGRHAGWIAL-----ESGMAAGA--HGICLERPPDP 198
QY 274 EBAKRLAEVMDK-----VDCVNI FVSGAGVEALVAEMQKQGEVPRDAFGIKLDAVN 328
DB 199 AELVWVVERFARFGKPAVLC-----VAEGA--HPVGSMDYSHGAI--DQGHRRFOGI- 249
QY 329 PKWFGGEPQOMIGAE--TLVQKSGYFARASAVNDMRLIKSCADLAVECAFRRREG 385
DB 250 -GTALAAELERLIGKEARPV--LGHVORGSTPTAVDVLATRFGMHVAEVAHRRGFG 304

```

```

QY 209 AENNANPRLIVHEVMGNCGLTAATAOEYKLLDRAEWLPELGLTRESYEVHAFVPE 268
DB 156 AESH---KRVIVCVWGRHGMIT-----YAGIAGAD-----VILVEE 192
QY 269 MAIDLEAKKL--REVMKVDVNIIVFVSGAGVEALVAMQAKGQVPRDAPGHTKIDA 326
DB 193 IPADLAKVAEHIOGRHAGRTFSI-VVAAGTRIKLSADQOEQLVTSGLADKGRDRLGG 251
QY 327 VNRKWFGEOPACQITGKTLVQKSGYFAPASASNDMDL-----IKSCADLAVCAR 381
DB 252 V--GTILAHETERGTPE-TRVSVLGHIOGGAFTADRLVATRYGVHAC-DVVARGEBG 307
QY 382 RESGVIGHD 390
DB 308 KMAALRGND 316

```

```

RESULT 12
US-10-369-493-2905
; Sequence 2905, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2905
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Thermotoga maritima
US-10-369-493-2905

```

```

Query Match 11.1%; Score 253; DB 12; Length 319;
Best Local Similarity 26.6%; Pred. No. 5.4e-17;
Matches 95; Conservative 60; Mismatches 126; Indels 76; Gaps 14;
QY 33 KQVAILTAGGLAPCLNSAIGSLIERYTEIDPSIEIICRYGKYKGLLGDSPVTAVERKK 92
DB 2 KKLAVLTSGGDAFGNNAVAV-RY-GVRQGLEVIQVRRGYSGLIDGF--VKLEVDY 57
QY 93 AGVLORRGSVIGNSRYKLTNVVDQYKSGLVKESGEPQKVAADQLVKQGVDLHTIGDD 152
DB 58 AGITER-GGITLTSTRKEEFT-----EEGRLEAKQIKQHIEGLVVIQEG 104
QY 153 TTTAADLAFLARNYGLTVIGLPTVNDVFPFKOSLGAWTAAEQGARFTMNVAAEN 212
DB 105 SLF-----GAILYEEHKIPVVGIPATINDIGLTMCGIVDTCLNTVWDANQKIDTAS 159
QY 213 ANRPMILIVHEVMGNCGLTAATAOEYKLLDRAEWLPELGLTRESYEVHAFVPEMAID 272
DB 160 SHRAFTV-EVMRSHSGYIALMG-----LVYGA-----AIVVEIPVD 198
QY 273 LEAEAKRLREV--MDKVDVNIIVFVSGAGVEALVAMQAKGQVPRDAPGHTKIDAVNP 330
DB 199 YSLADRLIEERRRGINSI-ITVAGASATVVA-----HLE----- 236
QY 331 KMGQOPAKITGKTLVQKSGYFAPASASNDMDLISKADLAVECAFRESGVI 387
DB 237 -----YRIGYE-TRITILGHVQSGSFPAFDRRLALSMGVEAVDALDDEVDM 284

```

RESULT 13  
US-10-369-493-9669

```

; Sequence 9669, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9669
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Desulfotobacterium hafnense
US-10-369-493-9669

```

```

Query Match 11.0%; Score 250.5; DB 12; Length 357;
Best Local Similarity 27.4%; Pred. No. 1.1e-16;
Matches 101; Conservative 63; Mismatches 136; Indels 69; Gaps 17;
QY 34 KVALITAGGLAPCLNSAIGSLIERYTEIDPSIEIICRYGKYKGLLGDSPVTAVERKK 93
DB 1 KVALITGGGCGPGLNATVRAVK--TACRGIVAVIOGFRFRAVAGDREPL--ELKQVS 56
QY 94 GVLORFGSVIG-----NSRVKLTNVKDCVKGQLVKESGEPQKVAADQLVKQGVDLHTI 148
DB 57 GILPR-GGITLTSTRNDPRAVPTQV-----GEHQVQDRSAEVLDRKAEIGDILAI 109
QY 149 GGDQVNTAADLAFLARNYGLTVIGLPTVNDVFPFKOSLGAWT--AAEQGARFTMNV 206
DB 110 GGGSLSTAEFA-----KQGLKAVGVRTINDLACTDITFGPOTAAATQDADLRH 163
QY 207 VVAENNANPRLIVHEVMGNCGLTAATAOEYKLLDRAEWLPELGLTRESYEVHAFV 266
DB 164 TTAESH---HRIMILEVMGRYAGMIAL-----YAGVAGAD-----VILI 200
QY 267 PENALIDE--AEAKRLREVMKVDVNIIVFVSGA---GVEALVAMQAKGQVPRDAPGH 321
DB 201 PELPYQLESIAEAVQRRARLGRFST-ITVAGAKPLGGDMV-ERTMSGRTDP----- 252
QY 322 IKLDVAVPQKWFGEQA---QMTGAEKTVQKSGYFAPASASNDMDLISKADLAVEC 378
DB 253 IKLGGI-----GAKLADLEKVTDMETVTVLGHLORGSGPLAVDRVLTSTRYGVAVEA 306
QY 379 AFRRESGVI 387
DB 307 ALAGDPGM 315

```

```

RESULT 14
US-10-369-493-21998
; Sequence 21998, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374

```

SEQ ID NO 21998  
 LENGTH: 987  
 TYPE: PR  
 ORGANISM: Saccharomyces cerevisiae  
 US-10-369-493-21998

Query Match 10.8%; Score 246.5; DB 12; Length 987;  
 Best Local Similarity 23.9%; Pred. No. 1.4e-15;  
 Matches 103; Conservative 68; Mismatches 135; Indels 125; Gaps 19;

QY 31 KKKKVALITAGGAPCLNSAIGSLIERYTEIDPSIEIICRGYKGLTSGSYVTAER 90  
 DB 204 KKKKJAVMTSGDSEPMNAAYAVR--TGIFGCDVAVTEGIEGLRGSKY-----LK 256  
 QY 91 KKA-----GVICRFQSVIGNSRYKLTNVKDCVKGGLVKEGEDPQKVAADQLVKDVIL 145  
 DB 257 KNAMEVVRGMLSE--GGTILGTA-----RSMFRRKRGSRQAGNLISQGDAL 303  
 QY 146 HTIGDDTNTAAADL-----AAPLARNY-----GLTVIGLPTVNDVPEPI 187  
 DB 304 VVCGGDSGLT--GADLFRHEWPSLVDELVAEGRFTEEVAPYKMLSTIVGLVGSIDNDMSGT 362  
 QY 188 KSLGAMTAEGQAFYFVAVVANNANPRMLIVHEVMGRNGMWT---AATAQERYKLT 243  
 DB 363 DSTIGAVSALERICEMWDYIDATKSHSRFAV--EVMGRHCGMLALMAGIATGADY----- 417  
 QY 244 DRAEMLPBILGLTBESYEVAHVPEMAIDLEAEAKRLREVMDKVCV---NIFVEGA 298  
 DB 418 -----LFIPEFAVPHGKMODEKWEVCORHRSKGRNNITIVAGA 457  
 QY 239 -----GVEAIVAMQAKQOEVRDAFGHKL--DAVNGKXWFGSEQFQAMCAEK 345  
 DB 458 LDDQNPVTANDVDALIEL---GLDTVTTLGHVQGGTRVADRLAT--LQGVDAVK 512  
 QY 346 TLVQSGYFAPASASN-----VDNRILKSCADLAWECAFRREGVIGHDS 391  
 DB 513 AVL-----FTPEYSPFLIGLEKNIIMPIVESVKLTSSVA-----TALEKDF 557  
 QY 392 DNGNVLRAIEF 402  
 DB 558 DKALSLRDER 568

RESULT 15  
 US-09-815-242-13785  
 Sequence 13785, Application US/09815242  
 Patent No. US20020061569A1

GENERAL INFORMATION:  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Karl L.  
 APPLICANT: Zysek, Judith W.  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John D.  
 APPLICANT: Carr, Grant J.  
 APPLICANT: Yamamoto, Robert T.  
 TITLE OF INVENTION: Identification of Essential Genes in  
 FILE REFERENCE: BLITRA.011A  
 CURRENT APPLICATION NUMBER: US/09/815.242  
 CURRENT FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 NUMBER OF SEQ ID NOS: 14110  
 SOFTWARE: FastSeq for windows version 4.0  
 SEQ ID NO 13785  
 LENGTH: 320  
 TYPE: PR  
 ORGANISM: Salmonella typhi  
 US-09-815-242-13785

Query Match 10.5%; Score 240.5; DB 9; Length 320;  
 Best Local Similarity 24.8%; Pred. No. 1e-15;  
 Matches 102; Conservative 59; Mismatches 122; Indels 129; Gaps 18;

QY 33 KKVAILTAGGAPCLNSAIGSLIERYTEIDPSIEIICRGYKGLTSGSYVTAER 87  
 DB 3 KKGIVLTSGLDAPGNMAIRGVYR--AALTSELEVMGTYDGLGIEYRMVQLDRYSVSD 60  
 QY 88 EVRKAGVLCRFQSVIGNSRYKLTNVKDCVKGGLVKEGEDPQKVAADQLVKDVILHT 147  
 DB 61 MNR-----GGTFLGSAR--PPEPRD-----ENIRAVAIENLKKRGIDALLV 100  
 QY 146 IGGDDTNTAAADLAAFLARNYGLTVIGLPTVNDVPEPIKOSIGAMTAEGQAFYFVAV 207  
 DB 101 IGGDSYKAKRL-----TEMGFPCIGLPGTIDNDIGCTYTTIGYFALGTYVEALDR 154  
 QY 208 VANNANPRMLIVHEVMGRNGMWTAAATQERYKLTDRAEWL--PELGLTBESYEVAHV 265  
 DB 155 RDTSSSHORISIV--EVMGRYCGDLTAA-----IAGCEPIVAVPEVEFNE----- 200  
 QY 266 VPEMAIDLEAEK-----RLREVMDKVCNIPFSEAGVE--AIVAMQAK 310  
 DB 201 -----DLVAETKGIAGKKAHVAITEHCVDDELAHFTEKGTGEBTATVIGHQORG 254  
 QY 311 GQEVPRDAFGHKLDAVNPQXWFGSQFQAMIGAETIYQKSGYFAPASASNVDMLIKS 370  
 DB 255 GSPVYD-----RILASRGAVA----- 272  
 QY 371 CADLAVE-----CAFRREGVIGHDENGNVLRAIEPPRIKGGKPPNIDTW 417  
 DB 273 -IDLLEHGGRGCGVIGNEQLVHED-----IIDAIENMK--RPFK--SDW 312

Search completed: January 29, 2004, 14:43:34  
 Job time : 40 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 15:54:19 : Search time 26.8953 Seconds  
(without alignments)  
3044.503 Million cell updates/sec

Title: US-09-941-947a-8  
Perfect score: 1985  
Sequence: 1 MKGICILGATGSGIVSTLDV.....LQADQAREVARDIKTLVA 394

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10C\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1985	100.0	394	10	US-09-934-903-4 Sequence 4, Appl1
2	1985	100.0	394	10	US-09-934-868-4 Sequence 64, Appl1
3	1985	100.0	394	11	US-09-941-947a-8 Sequence 8, Appl1
4	1215.5	61.2	396	12	US-10-381-779-117 Sequence 117, App
5	1052.5	53.0	394	12	US-10-381-779-119 Sequence 119, App
6	1045	52.6	398	12	US-10-381-779-100 Sequence 100, App
7	1045	52.6	398	12	US-10-381-779-118 Sequence 118, App
8	1042	52.5	405	12	US-10-381-779-122 Sequence 122, App
9	1007	50.7	397	12	US-10-381-779-99 Sequence 99, Appl1
10	1007	50.7	397	12	US-10-381-779-120 Sequence 120, App
11	999.5	50.4	388	12	US-10-381-779-101 Sequence 101, App
12	999.5	50.4	388	12	US-10-381-779-116 Sequence 116, App
13	994.5	50.1	386	12	US-10-381-779-97 Sequence 97, Appl1
14	892.5	45.0	394	12	US-10-381-779-103 Sequence 103, App
15	892.5	45.0	394	12	US-10-381-779-121 Sequence 121, App

15	886	44.6	402	12	US-10-381-779-102 Sequence 102, App
17	886	44.6	402	12	US-10-381-779-123 Sequence 123, App
18	818.5	41.2	477	9	US-09-923-556-6 Sequence 6, Appl1
19	818.5	41.2	477	10	US-09-967-025-2 Sequence 2, Appl1
20	818.5	41.2	477	12	US-10-381-779-128 Sequence 128, App
21	817.5	41.2	477	14	US-10-047-412a-10 Sequence 10, Appl1
22	814.5	41.0	477	9	US-09-923-556-2 Sequence 2, Appl1
23	813.5	41.0	477	14	US-10-047-412a-29 Sequence 29, Appl1
24	798.5	40.2	388	12	US-10-381-779-98 Sequence 98, Appl1
25	798.5	40.2	388	12	US-10-381-779-125 Sequence 125, App
26	731.5	36.9	385	12	US-10-128-713a-4 Sequence 4, Appl1
27	729.5	36.8	486	12	US-10-381-779-131 Sequence 131, App
28	721.5	36.3	387	12	US-10-259-194a-268 Sequence 268, App
29	718	36.2	386	12	US-10-381-779-124 Sequence 124, App
30	697	35.1	394	15	US-10-156-761-10100 Sequence 10100, A
31	682	34.9	392	10	US-09-738-626-5709 Sequence 5709, App
32	681	34.8	436	10	US-09-712-363-258 Sequence 258, App
33	691	34.8	436	12	US-10-381-779-104 Sequence 104, App
34	691	34.8	436	12	US-10-381-779-127 Sequence 127, App
35	681	34.3	406	12	US-10-381-779-126 Sequence 126, App
36	620	31.2	340	12	US-10-289-762-362 Sequence 362, App
37	596	30.0	1588	12	US-09-925-778-2 Sequence 2, Appl1
38	486	24.5	356	12	US-10-381-779-130 Sequence 130, App
39	473	23.8	368	12	US-09-882-227-500 Sequence 500, App
40	428.5	21.6	177	12	US-10-032-395-1 Sequence 1, Appl1
41	251	12.6	98	12	US-09-864-408a-8646 Sequence 8646, App
42	114	5.7	425	12	US-10-369-493-21773 Sequence 21773, A
43	113.5	5.7	1073	15	US-10-156-761-12156 Sequence 12156, A
44	109.5	5.5	6146	15	US-10-156-761-10436 Sequence 10436, A
45	109	5.5	426	12	US-10-389-493-21804 Sequence 21804, A

## ALIGNMENTS

RESULT 1  
US-09-934-903-4  
Sequence 4, Application US/09934903  
Patent No. US20020102690A1  
GENERAL INFORMATION:  
APPLICANT: Kofas, Matheos  
APPLICANT: Odem, J. Martin  
APPLICANT: Schenzle, Andreas J.  
APPLICANT: No. US20020102690A1ton, Kelley C.  
APPLICANT: Tomb, Jean-Francois  
APPLICANT: Rouviere, Pierre  
APPLICANT: Picataggio, Stephen  
APPLICANT: Cheng, Qiong  
TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production  
FILE REFERENCE: C11646 US NA  
CURRENT APPLICATION NUMBER: US/09/934,903  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/229,907  
PRIOR FILING DATE: September 1, 2001  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 4  
LENGTH: 394  
TYPE: PRT  
ORGANISM: Methylobionas 16a  
FEATURE:  
OTHER INFORMATION: Amino acid sequences encoded by ORF2  
US-09-934-903-4

Query Match 100.0%, Score 1985, DB 10, Length 394;  
Best Local Similarity 100.0%, Pred. No. 7.2e-186;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKGICILGATGSGIVSTLDVVAHPDKQVVALTANGNIDALYECIAHREYAVVWES 60  
Db 1 MKGICILGATGSGIVSTLDVVAHPDKQVVALTANGNIDALYECIAHREYAVVWES 60  
QY 61 KVAERKQRLAASPVAADIVLSSSEALQVAVLENDVYMAAIVGAAGLPTLAAKAGKT 120

Db 61 KVAEFKRIASPVADIKVLSGSEALQOVATLENDVTMAAIVGAAGLFTTLAAARXGT 120  
Qy 121 VLANKKALVMSGGIFMVAVSDSGAVLLPIDSSENAIFQCPAGYTPGHTAKARRILLT 180  
Db 121 VLANKKALVMSGGIFMVAVSDSGAVLLPIDSSENAIFQCPAGYTPGHTAKARRILLT 180  
Qy 181 ASGGFRRTPIETLSSTVPDOAVAHPRKDMGRKISVDSATMMKGLIELACLLFEMBED 240  
Db 181 ASGGFRRTPIETLSSTVPDOAVAHPRKDMGRKISVDSATMMKGLIELACLLFEMBED 240  
Qy 241 QIEVVIHQSIHSMVVDVDSVLAQKGNPDMRTPIAHAMAMPERFDSGVAPLDIEVGH 300  
Db 241 QIEVVIHQSIHSMVVDVDSVLAQKGNPDMRTPIAHAMAMPERFDSGVAPLDIEVGH 300  
Qy 301 MDEFKPDIKRPPCRLAYEALIKSGIMPTVLANAEIAVAFLNEEVEKFTDIIVIIERSM 360  
Db 301 MDEFKPDIKRPPCRLAYEALIKSGIMPTVLANAEIAVAFLNEEVEKFTDIIVIIERSM 360  
Qy 361 AOFKPDAGSLLEVLQADQDAREVARDIITLVA 394  
Db 361 AOFKPDAGSLLEVLQADQDAREVARDIITLVA 394

RESULT 2  
US-09-934-868-64  
Sequence 64, Application US/09934868  
Patent No. US20020137190A1  
GENERAL INFORMATION:  
APPLICANT: Koffas, Mattheos  
APPLICANT: Odom, James M.  
APPLICANT: Schenkel, Andreas J.  
TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN  
FILE REFERENCE: CL1596 US NA  
CURRENT APPLICATION NUMBER: US/09/934, 868  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/229, 858  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 64  
LENGTH: 394  
TYPE: PRT  
ORGANISM: Methylomonas 16a  
FEATURE:  
OTHER INFORMATION: Amino acid sequences encoded by DXR  
US-09-934-868-64

Query Match 100.0%; Score 1985; DB 10; Length 394;  
Best Local Similarity 100.0%; Pred. No. 7, 2e-186;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKGICILGATSGISGVSTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHHEPVAVVMES 60  
Db 1 MKGICILGATSGISGVSTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHHEPVAVVMES 60  
Qy 61 KVAEFKRIASPVADIKVLSGSEALQOVATLENDVTMAAIVGAAGLFTTLAAARXGT 120  
Db 61 KVAEFKRIASPVADIKVLSGSEALQOVATLENDVTMAAIVGAAGLFTTLAAARXGT 120  
Qy 121 VLANKKALVMSGGIFMVAVSDSGAVLLPIDSSENAIFQCPAGYTPGHTAKARRILLT 180  
Db 121 VLANKKALVMSGGIFMVAVSDSGAVLLPIDSSENAIFQCPAGYTPGHTAKARRILLT 180  
Qy 181 ASGGFRRTPIETLSSTVPDOAVAHPRKDMGRKISVDSATMMKGLIELACLLFEMBED 240  
Db 181 ASGGFRRTPIETLSSTVPDOAVAHPRKDMGRKISVDSATMMKGLIELACLLFEMBED 240  
Qy 241 QIEVVIHQSIHSMVVDVDSVLAQKGNPDMRTPIAHAMAMPERFDSGVAPLDIEVGH 300  
Db 241 QIEVVIHQSIHSMVVDVDSVLAQKGNPDMRTPIAHAMAMPERFDSGVAPLDIEVGH 300  
Qy 301 MDEFKPDIKRPPCRLAYEALIKSGIMPTVLANAEIAVAFLNEEVEKFTDIIVIIERSM 360  
Db 301 MDEFKPDIKRPPCRLAYEALIKSGIMPTVLANAEIAVAFLNEEVEKFTDIIVIIERSM 360

Db 301 MDEFKPDIKRPPCRLAYEALIKSGIMPTVLANAEIAVAFLNEEVEKFTDIIVIIERSM 360  
Qy 361 AOFKPDAGSLLEVLQADQDAREVARDIITLVA 394  
Db 361 AOFKPDAGSLLEVLQADQDAREVARDIITLVA 394

RESULT 3  
US-09-941-947a-8  
Sequence 8, Application US/09941947A  
Publication No. US20030003528A1  
GENERAL INFORMATION:  
APPLICANT: Bzostowicz, Patricia C.  
APPLICANT: Cheng, Qiong  
APPLICANT: DiCosimo, Deana J.  
APPLICANT: Koffas, Mattheos  
APPLICANT: Miller, Edward S. Jr.  
APPLICANT: Odom, J. Martin  
APPLICANT: Picatagallo, Steve  
APPLICANT: Rouviere, Pierre E.  
TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE  
FILE REFERENCE: CL1903 US NA  
CURRENT APPLICATION NUMBER: US/09/941, 947A  
CURRENT FILING DATE: 2001-09-01  
PRIOR APPLICATION NUMBER: 60/229, 907  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229, 858  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 60  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 8  
LENGTH: 394  
TYPE: PRT  
ORGANISM: Methylomonas 16a  
US-09-941-947a-8

Query Match 100.0%; Score 1985; DB 11; Length 394;  
Best Local Similarity 100.0%; Pred. No. 7, 2e-186;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKGICILGATSGISGVSTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHHEPVAVVMES 60  
Db 1 MKGICILGATSGISGVSTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHHEPVAVVMES 60  
Qy 61 KVAEFKRIASPVADIKVLSGSEALQOVATLENDVTMAAIVGAAGLFTTLAAARXGT 120  
Db 61 KVAEFKRIASPVADIKVLSGSEALQOVATLENDVTMAAIVGAAGLFTTLAAARXGT 120  
Qy 121 VLANKKALVMSGGIFMVAVSDSGAVLLPIDSSENAIFQCPAGYTPGHTAKARRILLT 180  
Db 121 VLANKKALVMSGGIFMVAVSDSGAVLLPIDSSENAIFQCPAGYTPGHTAKARRILLT 180  
Qy 181 ASGGFRRTPIETLSSTVPDOAVAHPRKDMGRKISVDSATMMKGLIELACLLFEMBED 240  
Db 181 ASGGFRRTPIETLSSTVPDOAVAHPRKDMGRKISVDSATMMKGLIELACLLFEMBED 240  
Qy 241 QIEVVIHQSIHSMVVDVDSVLAQKGNPDMRTPIAHAMAMPERFDSGVAPLDIEVGH 300  
Db 241 QIEVVIHQSIHSMVVDVDSVLAQKGNPDMRTPIAHAMAMPERFDSGVAPLDIEVGH 300  
Qy 301 MDEFKPDIKRPPCRLAYEALIKSGIMPTVLANAEIAVAFLNEEVEKFTDIIVIIERSM 360  
Db 301 MDEFKPDIKRPPCRLAYEALIKSGIMPTVLANAEIAVAFLNEEVEKFTDIIVIIERSM 360  
Qy 361 AOFKPDAGSLLEVLQADQDAREVARDIITLVA 394  
Db 361 AOFKPDAGSLLEVLQADQDAREVARDIITLVA 394

RESULT 4  
US-10-381-779-117  
Sequence 117, Application US/10381779

```

; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jensen, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/002US1
; CURRENT APPLICATION NUMBER: US/10/381,779
; PRIOR FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-381-779-117

Query Match      61.2%; Score 1215.5; DB 12; Length 396;
Best Local Similarity 64.1%; Pred. No. 2.1e-110;
Matches 245; Conservative 48; Mismatches 88; Indels 1; Gaps 1;

QY 4 ICILGATGSIQVSTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHPEYAVVVMESKVA 63
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 7 ISVIGATGSIQSTLDVVAQHPDRYEAFALTGFSRLAELEALCLRHRPVAVVPEQAAAI 66
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 64 EFKORIAAPVADIKVLSGSEALQOVATLNAVDTVMVAIVGAAGLPLTLAAAKGKTVLL 123
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 67 ALQSLAAAGIR-TRVLFGEQALCEVASAPEVDVMVAALVGAAGLPLTLAAVNGKRVLL 125
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 124 ANKEALVMSGOIFMQAVSDSGAVLLPIDSEHNAIPQCMFAGYTPGHTAKQARRILLTASG 183
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 126 ANKEALVMSGALFMQAVKSGAVLLPIDSEHNAIPQSLPRVYADGLERGVGRILLTASG 185
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 184 GPFRTPIETLSSVTPDQAVAHPEKMDGKRTSVDSATMMNGKLELIEACLLFNMEDPOIE 243
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 GPFRTPIETLSSVTPDQAVAHPEKMDGKRTSVDSATMMNGKLELIEACLLFNMEDPOIE 245
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 244 VVIHPQSIHSMVDVYDGSVLAQMGNDPMDRTPIAHANAMPERFDGVAPLDIFEVGHMDF 303
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 246 VVIHPQSIHSMVDVYDGSVLAQMGNDPMDRTPIAHANAMPERFDGVAPLDIFEVGHMDF 305
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 304 EKPLKRPFCRLAIVEALKSGGIMPTVLANAETAVAFLENEVKFTDIATITERSMAQF 363
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 306 QRPEDRPFCLRLAIVEALKSGGIMPTVLANAETAVAFLENEVKFTDIATITERSMAQF 365
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 364 KPDAGSLELVLAQDQDAREVARDIITKL 392
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 366 AVTAVESLDQVLAADRRARSRVA 387
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
US-10-381-779-119
; Sequence 119, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jensen, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/002US1
; CURRENT APPLICATION NUMBER: US/10/381,779
; PRIOR FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119
; SSEQ ID NO 119

```

```

; LENGTH: 394
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-381-779-119

Query Match      53.0%; Score 1052.5; DB 12; Length 394;
Best Local Similarity 55.5%; Pred. No. 2.1e-94;
Matches 216; Conservative 49; Mismatches 123; Indels 1; Gaps 1;

QY 4 ICILGATGSIQVSTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHPEYAVVVMESKVA 63
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 6 LITLSTGSIQSTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHPEYAVVVMESKVA 65
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 64 EFKORIAAPVADIKVLSGSEALQOVATLNAVDTVMVAIVGAAGLPLTLAAAKGKTVLL 123
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 66 RLEALLKEDGTR-TRVLFGEQALCEVASAPEVDVMVAALVGAAGLPLTLAAVNGKRVLL 124
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 124 ANKEALVMSGOIFMQAVSDSGAVLLPIDSEHNAIPQCMFAGYTPGHTAKQARRILLTASG 183
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 126 ANKEALVMSGALFMQAVKSGAVLLPIDSEHNAIPQSLPRVYADGLERGVGRILLTASG 184
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 184 GPFRTPIETLSSVTPDQAVAHPEKMDGKRTSVDSATMMNGKLELIEACLLFNMEDPOIE 243
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 GPFRTPIETLSSVTPDQAVAHPEKMDGKRTSVDSATMMNGKLELIEACLLFNMEDPOIE 244
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 244 VVIHPQSIHSMVDVYDGSVLAQMGNDPMDRTPIAHANAMPERFDGVAPLDIFEVGHMDF 303
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 246 VVIHPQSIHSMVDVYDGSVLAQMGNDPMDRTPIAHANAMPERFDGVAPLDIFEVGHMDF 304
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 304 EKPLKRPFCRLAIVEALKSGGIMPTVLANAETAVAFLENEVKFTDIATITERSMAQF 363
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 306 QRPEDRPFCLRLAIVEALKSGGIMPTVLANAETAVAFLENEVKFTDIATITERSMAQF 364
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 364 KPDAGSLELVLAQDQDAREVARDIITKL 392
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 366 AVTAVESLDQVLAADRRARSRVA 387
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
US-10-381-779-100
; Sequence 100, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jensen, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/002US1
; CURRENT APPLICATION NUMBER: US/10/381,779
; PRIOR FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 100
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-381-779-100

Query Match      52.6%; Score 1045; DB 12; Length 398;
Best Local Similarity 54.9%; Pred. No. 1.1e-93;
Matches 219; Conservative 56; Mismatches 118; Indels 6; Gaps 3;

QY 1 MKGICILGATGSIQVSTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHPEYAVVVMESKVA 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MKQITLIGSTGSIQSTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHPEYAVVVMESKVA 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 KVAEFKORIAAPVADIKVLSGSEALQOVATLNAVDTVMVAIVGAAGLPLTLAAAKGKTVLL 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 S-AKGLKTMKQQGGSRTVLSGGQAAADVAEDVDQVVAIVGAAGLPLTLAAAKGKTVLL 119
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Qy 121 VLANKKALVMSGOIFMOAVSDSGAVLPIIDSEHNALFOCM--PAGYTPGHAKQAR--- 175
Db 120 ILANKESLVTGCRMLMDAVKQSKAGLLPVDSEHNALFQSLPQPIQHNLGVADLEQNGV 179
Qy 176 RLLLTASGGPFRRTPIETLSVTPDQAVAHPRKMDGRKISVDSATMANKGELIIPACLLF 235
Db 180 SILLTSGGPFRETPRLDATTMTPDQACHFPMWSMKRKISVSATMANKGLEYIARWL 239
Qy 236 NMEPDQLEVIHPOSIHSMVDYVDSVLAQMGNDMPRTPIAHAMAMPERFDSGVAPLDI 295
Db 240 NASASQMEVLHPOSIVHSVTRQODSVLAQLOEPDKRTPIAHAMAMPERVNSGVKPLDF 299
Qy 296 FEVGHMDFEKDPDKRPPCLRLAYEALIKSGGIMPTVLNANEIAYEAFLEBEVKFTDIAT 355
Db 300 CKLSALTFPAADRYECLKLAMEAFEGQATTAALNANEITVAFAAQOIRFTDIAL 359
Qy 356 IERSMAQFKPDAGSLELVLAQADQAREVARDIITLVA 394
Db 360 NLSVLERMDREBPQCVDDVLSDANAREVAREVRLAS 398

RESULT 7
US-10-381-779-118
; Sequence 118, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jessem, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/002051
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-381-779-118

Query Match 52.6%; Score 1045; DB 12; Length 398;
Best Local Similarity 54.9%; Pred. No. 1,1e-93;
Matches 219; Conservative 56; Mismatches 118; Indels 6; Gaps 3;

Qy 1 MKGICILGATGSGTLDVVAHPPKQYVALTANGNDIALYEQCLAHHPREYAVVWMS 60
Db 1 MKQTLTIGSTIGSTLTDVVRNHPHRVVALVAGNVTRMVEQCLERSPRAYWMDDA 60
Qy 61 KVAEFKQRIASPVADIKVLSGSEALQOVATLENVTVMALYGAAGLLPTLAAAGKT 120
Db 61 S-AKLTKMLQOQSRTREVLSGQOACDMALBVDVMAIVGAAGLLPTLAAIAGKT 119
Qy 121 VLANKKALVMSGOIFMOAVSDSGAVLPIIDSEHNALFOCM--PAGYTPGHAKQAR--- 175
Db 120 ILANKESLVTGCRMLMDAVKQSKAGLLPVDSEHNALFQSLPQPIQHNLGVADLEQNGV 179
Qy 176 RLLLTASGGPFRRTPIETLSVTPDQAVAHPRKMDGRKISVDSATMANKGELIIPACLLF 235
Db 180 SILLTSGGPFRETPRLDATTMTPDQACHFPMWSMKRKISVSATMANKGLEYIARWL 239
Qy 236 NMEPDQLEVIHPOSIHSMVDYVDSVLAQMGNDMPRTPIAHAMAMPERFDSGVAPLDI 295
Db 240 NASASQMEVLHPOSIVHSVTRQODSVLAQLOEPDKRTPIAHAMAMPERVNSGVKPLDF 299
Qy 296 FEVGHMDFEKDPDKRPPCLRLAYEALIKSGGIMPTVLNANEIAYEAFLEBEVKFTDIAT 355
Db 300 CKLSALTFPAADRYECLKLAMEAFEGQATTAALNANEITVAFAAQOIRFTDIAL 359

```

```

Qy 356 IERSMAQFKPDAGSLELVLAQADQAREVARDIITLVA 394
Db 360 NLSVLERMDREBPQCVDDVLSDANAREVAREVRLAS 398

RESULT 8
US-10-381-779-122
; Sequence 122, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jessem, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/002051
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-381-779-122

Query Match 52.5%; Score 1042; DB 12; Length 405;
Best Local Similarity 53.2%; Pred. No. 2,3e-93;
Matches 210; Conservative 68; Mismatches 109; Indels 8; Gaps 3;

Qy 1 MKGICILGATGSGTLDVVAHPPKQYVALTANGNDIALYEQCLAHHPREYAVVWMS 60
Db 7 MKQTLTIGSTIGSTLTDVVRNHPHRVVALVAGNVTRMVEQCLERSPRAYWMDDA 66
Qy 61 KVAEFKQRIASPVADIKVLSGSEALQOVATLENVTVMALYGAAGLLPTLAAAGKT 120
Db 67 AKKALAKLKHQ-SQTTVLAQOQALCELAHPEALMVMAIVGAAGLLPTLSAVAGKR 125
Qy 121 VLANKKALVMSGOIFMOAVSDSGAVLPIIDSEHNALFOCM-----AGYTPGHAKQAR 174
Db 126 VLANKKALVTCGLFTDAVRESOQQLPVDSEHNALFQSLPEAQROIGFC-LSBLGI 184
Qy 176 RLLLTASGGPFRRTPIETLSVTPDQAVAHPRKMDGRKISVDSATMANKGELIIPACLL 234
Db 185 SKVLITSGGPFRTPIEBQEQITTPAQAVAHPRKMDGRKISVDSATMANKGLEYIARWL 244
Qy 236 NMEPDQLEVIHPOSIHSMVDYVDSVLAQMGNDMPRTPIAHAMAMPERFDSGVAPLDI 294
Db 245 FNAEAEAEVLIHPOSIHSMVTRIDSVIAQMGNDMPRTPIAHAMAMPERVNSGVKPLDF 304
Qy 296 FEVGHMDFEKDPDKRPPCLRLAYEALIKSGGIMPTVLNANEIAYEAFLEBEVKFTDIAT 354
Db 305 FYQANGLTFTPEQRPCLKLALDAFSAQOYATTMNANAEIYAVSFLDNKIKFTDIAR 364
Qy 356 IERSMAQFKPDAGSLELVLAQADQAREVARDIITLVA 389
Db 365 LNLVSVSLQPKIHICIEDVLEVDKARBELSOSII 399

RESULT 9
US-10-381-779-99
; Sequence 99, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jessem, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/002051

```

```

; CURRENT APPLICATION NUMBER: US/10/381,779
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-381-779-99

```

```

Query Match      50.7%; Score 1007; DB 12; Length 397;
Best Local Similarity 50.6%; Pred. No. 6e-90;
Matches 200; Conservative 76; Mismatches 111; Indels 8; Gaps 3;

```

```

QY 2 KGCITLGGATGSGVSTLTVVAHAPKQVVALTANGNIDALYEQCLAHHPYAVVMESK 61
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 4 QNIVLISGSGIGKSTLTVIENNPQKHAFLVGGKVEAMFECQIKFRPHFALDDVNA 63
QY 62 VAEFKORIAASPVADIKVLSGSEALQOVAATLENDVTMAAIVGAAGLLPTLAAKAGKT 121
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 64 AKILREKLIANH1-PTVLAGRAICELAAHPDADQIMASIVGAAGLLPTLSAVAGKRV 122
QY 122 LLAKEALVMSGOITMVAUSDGAVLLPDISEHNAIFQCMF-----AGTTPGHATAQAR 175
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 123 LLANKESLVTGQGLFDVAVKNVSKLTPVDSHNAIFQSLPPEAKKIGFCP-LSBLGVS 181
QY 176 RLLTLASGPPFRRTPIETLSVTPDQAVAHKPMDEKRIKISVDSATMNNKGLIELACILF 235
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 182 KILLGSGPPFRYTPLEQPTNITPQAVAHNWSGKKISVDSATMNNKGLIELYEARLFL 241
QY 236 NMEPDQIEVVIHPQSIHSMWDYVDGSLAQMGNPDMRTPIAHAMWPERDSGVAPLDI 295
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 242 NASAEEMEVIIHPQSIHSMWRYVDGSLVITQMGNDMRTPIAETMAVPHRTFAGVEPLDF 301
QY 296 FEVGMDEPEKPDLPKPCPLRLAYRAIKSGGIMPTVLANAEIAVEAFINBEVFTDIAYI 355
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 302 FKIKELTIEPDPNRYPLKALIDPAAGQVATTAMNANAEIAVQAFIDRQIGFMDIAXI 361
QY 356 IERSVAQFPDDAGSLVYLQADDPAREVARDIIR 390
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 362 NSKITERISPYTIQINIDVLEIDQAREIAKTLLR 396

```

```

RESULT 10
US-10-381-779-120
; Sequence 120, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jesseen, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/002US1
; CURRENT APPLICATION NUMBER: US/10/381,779
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 120
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-381-779-120

```

```

Query Match      50.7%; Score 1007; DB 12; Length 397;
Best Local Similarity 50.6%; Pred. No. 6e-90;

```

```

Matches 200; Conservative 76; Mismatches 111; Indels 8; Gaps 3;
QY 2 KGCITLGGATGSGVSTLTVVAHAPKQVVALTANGNIDALYEQCLAHHPYAVVMESK 61
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 4 QNIVLISGSGIGKSTLTVIENNPQKHAFLVGGKVEAMFECQIKFRPHFALDDVNA 63
QY 62 VAEFKORIAASPVADIKVLSGSEALQOVAATLENDVTMAAIVGAAGLLPTLAAKAGKT 121
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 64 AKILREKLIANH1-PTVLAGRAICELAAHPDADQIMASIVGAAGLLPTLSAVAGKRV 122
QY 122 LLAKEALVMSGOITMVAUSDGAVLLPDISEHNAIFQCMF-----AGTTPGHATAQAR 175
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 123 LLANKESLVTGQGLFDVAVKNVSKLTPVDSHNAIFQSLPPEAKKIGFCP-LSBLGVS 181
QY 176 RLLTLASGPPFRRTPIETLSVTPDQAVAHKPMDEKRIKISVDSATMNNKGLIELACILF 235
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 182 KILLGSGPPFRYTPLEQPTNITPQAVAHNWSGKKISVDSATMNNKGLIELYEARLFL 241
QY 236 NMEPDQIEVVIHPQSIHSMWDYVDGSLAQMGNPDMRTPIAHAMWPERDSGVAPLDI 295
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 242 NASAEEMEVIIHPQSIHSMWRYVDGSLVITQMGNDMRTPIAETMAVPHRTFAGVEPLDF 301
QY 296 FEVGMDEPEKPDLPKPCPLRLAYRAIKSGGIMPTVLANAEIAVEAFINBEVFTDIAYI 355
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 302 FKIKELTIEPDPNRYPLKALIDPAAGQVATTAMNANAEIAVQAFIDRQIGFMDIAXI 361
QY 356 IERSVAQFPDDAGSLVYLQADDPAREVARDIIR 390
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 362 NSKITERISPYTIQINIDVLEIDQAREIAKTLLR 396

```

```

RESULT 11
US-10-381-779-101
; Sequence 101, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jesseen, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/002US1
; CURRENT APPLICATION NUMBER: US/10/381,779
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Zymomonas mobilis
US-10-381-779-101

```

```

Query Match      50.4%; Score 999.5; DB 12; Length 388;
Best Local Similarity 50.9%; Pred. No. 3.2e-89;
Matches 199; Conservative 74; Mismatches 109; Indels 9; Gaps 3;
QY 4 ICIGATGSGVSTLTVVAHAPKQVVALTANGNIDALYEQCLAHHPYAVVMESKVA 63
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 7 VIVLAGATGSGIGKSTLTVIENNPQKHAFLVGGKVEAMFECQIKFRPHFALDDVNA 66
QY 64 VAEFKORIAASPVADIKVLSGSEALQOVAATLENDVTMAAIVGAAGLLPTLAAKAGKT 123
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 67 DLKEALAGS---SVEAAGADALVEAMM-GADVTMAAIGCAGLKATLAIRGKVAL 122
QY 124 LLAKEALVMSGOITMVAUSDGAVLLPDISEHNAIFQCMGAGTTPGHATAQAR 183
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 123 AKRESLVASGGLMDAVAREHGTLLPVDSEHNAIFQCF-----PHNRDYYRKLITIASG 177
QY 184 GPFRTPIETLSVTPDQAVAHKPMDEKRIKISVDSATMNNKGLIELACILF 243
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```



Db 178 GFFRTSLAEMACTPERAVOHENMSGAKISIDATMKNKLELEAVHLFOIPEKEE 237  
 Qy 244 VVIHPOSIHSMVDYDGSVLAQMGNDPKRPIAHMAMPERFDGVAFLDIEVGERDF 303  
 Db 238 IIVHPOSVIHSWVEIDGSLIAQISDPKRTPIGHTLAMPKEMETPAESLDFTKRODWF 297  
 Qy 304 EKEDLKRFPCILRLAYEAIKSGIMPTVLAANEIYAEALNEVEKFTDIIVIERSMQF 363  
 Db 298 EADPYERFPALTLAMESIKSGGAPPAWMAANEIYAAFLDKKIGFTDIKIVERTLDHY 357  
 Qy 364 KPDDAGSLLEVLQADQDAREVARDIKITIVA 394  
 Db 358 TPATPSSLEDVPAIDNEARQAALAMESLPA 388

RESULT 12  
 US-10-381-779-116  
 ; Sequence 116, Application US/10381779  
 ; Publication No. US20030219798A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gokarn, Ravi R  
 ; APPLICANT: Jessen, Holly Jo  
 ; APPLICANT: Zidwick, Mary Jo  
 ; TITLE OF INVENTION: Isoprenoid Production  
 ; FILE REFERENCE: 12904/002US1  
 ; CURRENT APPLICATION NUMBER: US/10/381,779  
 ; CURRENT FILING DATE: 2003-03-28  
 ; PRIOR APPLICATION NUMBER: PCT/US01/30328  
 ; PRIOR FILING DATE: 2001-09-28  
 ; PRIOR APPLICATION NUMBER: 60/236,580  
 ; PRIOR FILING DATE: 2000-09-29  
 ; NUMBER OF SEQ ID NOS: 190  
 ; SOFTWARE: PatsSeq for Windows Version 4.0  
 ; SEQ ID NO 116  
 ; LENGTH: 388  
 ; TYPE: PRT  
 ; ORGANISM: Zymomonas mobilis  
 US-10-381-779-116

Query Match 50.4%; Score 999.5; DB 12; Length 388;  
 Best Local Similarity 50.9%; Pred. No. 3.2e-89;  
 Matches 199; Conservative 74; Mismatches 109; Indels 9; Gaps 3;  
 Qy 4 ICIIIGATGSGVSTLDVVAHPPDKYQVVALTANGNIDALYEQCLAHHPYAVVMESKVA 63  
 Db 7 VIVIGATGSGIHTLIDLERLDRYQVALTANGNVMDLADAKRTVAKKAVIADSLVN 66  
 Qy 64 EFFKRIASPVAIDIKVSGSEALQOVATLENDVTWMAIYGAAGLPTLAAAKAGKTVLL 123  
 Db 67 DLKEALGSG--SYEAAGADALVEAMM--GADWTMAIIGCAELKATTLAIAIRKKTVAL 122  
 Qy 124 ANKEALVMSGOIFQAVSDSGAVLLPTDSEHNAIFQCPHAGYTPGHTAXOARRILLTASG 183  
 Db 123 ANKESLVSAGGLMDIYAEHGTTLTPVSEHNAIFQCF-----PHNRDYYRRIIITASG 177  
 Qy 184 GFFRTPIETLSVTPDQAVAHPPKMDGRKISVDSATMKNKLELEIACLLFMEPEQIE 243  
 Db 178 GFFRTSLAEMACTPERAVOHENMSGAKISIDATMKNKLELEIYVHLFOIPEKEE 237  
 Qy 244 VVIHPOSIHSMVDYDGSVLAQMGNDPKRPIAHMAMPERFDGVAFLDIEVGERDF 303  
 Db 238 IIVHPOSVIHSWVEIDGSLIAQISDPKRTPIGHTLAMPKEMETPAESLDFTKRODWF 297  
 Qy 304 EKEDLKRFPCILRLAYEAIKSGIMPTVLAANEIYAEALNEVEKFTDIIVIERSMQF 363  
 Db 298 EADPYERFPALTLAMESIKSGGAPPAWMAANEIYAAFLDKKIGFTDIKIVERTLDHY 357  
 Qy 364 KPDDAGSLLEVLQADQDAREVARDIKITIVA 394  
 Db 358 TPATPSSLEDVPAIDNEARQAALAMESLPA 388

RESULT 13

US-10-381-779-97  
 ; Sequence 97, Application US/10381779  
 ; Publication No. US20030219798A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gokarn, Ravi R  
 ; APPLICANT: Jessen, Holly Jo  
 ; APPLICANT: Zidwick, Mary Jo  
 ; TITLE OF INVENTION: Isoprenoid Production  
 ; FILE REFERENCE: 12904/002US1  
 ; CURRENT APPLICATION NUMBER: US/10/381,779  
 ; CURRENT FILING DATE: 2003-03-28  
 ; PRIOR APPLICATION NUMBER: PCT/US01/30328  
 ; PRIOR FILING DATE: 2001-09-28  
 ; PRIOR APPLICATION NUMBER: 60/236,580  
 ; PRIOR FILING DATE: 2000-09-29  
 ; NUMBER OF SEQ ID NOS: 190  
 ; SOFTWARE: PatsSeq for Windows Version 4.0  
 ; SEQ ID NO 97  
 ; LENGTH: 386  
 ; TYPE: PRT  
 ; ORGANISM: Spilngomonas truegeri  
 US-10-381-779-97

Query Match 50.1%; Score 994.5; DB 12; Length 386;  
 Best Local Similarity 51.0%; Pred. No. 9.7e-89;  
 Matches 203; Conservative 69; Mismatches 109; Indels 17; Gaps 4;  
 Qy 1 MKGICIGATGSGVSTLDVVAHPPDKYQVVALTANGNIDALYEQCLAHHPYAVVMES 60  
 Db 2 VIVIGATGSGIHTLIDLERLDRYQVALTANGNVMDLADAKRTVAKKAVIADSLVN 66  
 Qy 61 KVAEFKRIASPVAIDIKVSGSEALQOVATLENDVTWMAIYGAAGLPTLAAAKAGKTV 120  
 Db 62 CLPALQERLAGS--GYEAAGADALVEAMM--GADWTMAIIGCAELKATTLAIAIRKKTVAL 117  
 Qy 121 VIANKALVMSGOIFQAVSDSGAVLLPTDSEHNAIFQCPHAGYTPGHTAXOARRILLTASG 176  
 Db 118 VALNKESLVSAGGLMDIYAEHGTTLTPVSEHNAIFQCF-----PHNRDYYRRIIITASG 177  
 Qy 177 ILTASGGFFRPIETLSVTPDQAVAHPPKMDGRKISVDSATMKNKLELEIACLLFMEPEQIE 243  
 Db 169 IIVHPOSVIHSWVEIDGSLIAQISDPKRTPIGHTLAMPKEMETPAESLDFTKRODWF 297  
 Qy 237 MEPEQIEVVIHPOSIHSMVDYDGSVLAQMGNDPKRPIAHMAMPERFDGVAFLDIEVGERDF 303  
 Db 229 VAAQGLAVLHRSQVHSWVEIDGSLIAQISDPKRTPIGHTLAMPKEMETPAESLDFTKRODWF 297  
 Qy 297 EYGHMDFEKPDLKEFPCLRLAYEAIKSGIMPTVLAANEIYAEALNEVEKFTDIIVIERSMQF 363  
 Db 289 TVGKLEFPENPDLRFPALTLAMESIKSGGAPPAWMAANEIYAAFLDKKIGFTDIKIVERTLDHY 357  
 Qy 357 ERSMQFEPDDAGSLLEVLQADQDAREVARDIKITIVA 394  
 Db 349 ADTLERYDPAPBTIDAVLAIDAEARLYAARVQDCA 386

RESULT 14  
 US-10-381-779-103  
 ; Sequence 103, Application US/10381779  
 ; Publication No. US20030219798A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gokarn, Ravi R  
 ; APPLICANT: Jessen, Holly Jo  
 ; APPLICANT: Zidwick, Mary Jo  
 ; TITLE OF INVENTION: Isoprenoid Production  
 ; FILE REFERENCE: 12904/002US1  
 ; CURRENT APPLICATION NUMBER: US/10/381,779  
 ; CURRENT FILING DATE: 2003-03-28  
 ; PRIOR APPLICATION NUMBER: PCT/US01/30328  
 ; PRIOR FILING DATE: 2001-09-28  
 ; PRIOR APPLICATION NUMBER: 60/236,580  
 ; PRIOR FILING DATE: 2000-09-29  
 ; NUMBER OF SEQ ID NOS: 190

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 103
; LENGTH: 394
; TYPE: PRN
; ORGANISM: Synechocystis sp. PCC 6803
US-10-381-779-103

```

Query Match	45.0%;	Score 892.5;	DB 12;	Length 394;
Best Local Similarity	47.3%;	Pred. No. 1e-78;		
Matches 183;	Conservative 63;	Mismatches 128;	Indels 13;	Gaps 3

```

QY      1 MKGCIIGAGSGSGVSTLELVAVARHPRTQVWALTTANGINDALYEGLAHBEVAVVMES 60
Db      2 VKRSLISUGSGSGISGTQLDI VTHHPAPFOVGLAAGNVALLAQVAFREIYAIRIOAE 61
QY      61 KVAEFKORIAASPVAIDIKVLSSGEALQQVATLENTVMAALVGAAGLLPTLAAKAKGT 120
Db      62 KLDELKAVALDELTDYGEKMYVAGEEVAVARVGDABSVTGVIGCAGLLPTMAALAAKGD 121
QY      121 VILANKAALVMSQGIKQKQANVSDSGANVLTIBSEHNATQCM---PAGTTPGHTAKQARI 177
Db      122 IALANKETILAGAPVVLPLVEKMGVLTLPADSESHAIKQCLQGVPEG-----GLRRI 173
QY      178 LLETPASGPPEPRTLETLSSTPDAQAAVAPXOMVGSKISVDSATVMNKKGLJLEACLEFNM 237
Db      174 ILTPASGAFPLDVEERLPRVTVQDALKEHNSXGQKRTTIDSKTILNKKGLJVAIHAHYFLGL 233
QY      238 EPDQIEVVIHPQSI IHSMTVDYDGSVLAQWENPDMRFTLAHAMAPFERDGSVAPDLIFB 297
Db      234 DYDHIDIVIHQGITIHSLEIYQDVTSLAQIGMPDRMLPLIYALASMPERITYDMWRPDLVYK 299
QY      298 VGHMDEKPELKEFPCLRLALYEAIKSGGIMPTVYLAANBLEAVEALNBSVKTIDANTIE 357
Db      294 AGSISFEEBPHDKYPCQQLAYGAGRGGAMPAYLANAEOAVALFLQKXISFLDIPRLIE 353
QY      358 RSMAGQFKRDPDAGS--LEIVLQADQDAR 382
Db      354 KTCDLVYGQNTASPDLDSTILAADQYMR 380

```

```

RESULT 15
US-10-381-779-121
: Sequence 121. Application US/10381779
: Publication No. US20030219798A1
: GENERAL INFORMATION:
: APPLICANT: Gokarn, Ravi R
: APPLICANT: Jessen, Holly
: APPLICANT: Zidwick, Mary Jo
: TITLE OF INVENTION: Isoprenoid Production
: FILE REFERENCE: 12904/002US1
: CURRENT APPLICATION NUMBER: US/10/381, 779
: CURRENT FILING DATE: 2003-03-28
: PRIOR APPLICATION NUMBER: PCT/US01/30328
: PRIOR FILING DATE: 2001-09-28
: PRIOR APPLICATION NUMBER: 60/236,580
: PRIOR FILING DATE: 2000-09-29
: NUMBER OF SEQ ID NOS: 190
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 121
: LENGTH: 394
: TYPE: PRT
: ORGANISM: Synechocystis sp. PCC 6803
US-10-381-779-121

```

Query Match	45.0%;	Score 892.5;	DB 12;	Length 394;
Best Local Similarity	47.3%;	Pred. No. 1e-78;		
Matches 183;	Conservative 63;	Mismatches 128;	Indels 13;	Gaps 3;

61 KVAEFGRIATSFVADIKVLSGSEALQVVATLENDVTYMAAIVGAAGLLPTLAARAKGT 120

Db	62	KEDLQAABELTDVQPMYVGESEBVENAVYGALESYVITIGVACAGLLPTMALIAGKD	121
Qy	121	VLLANKEALWMSGQIFMAQVSDSGAVLLPDISSENAIFQCK---	PAGYTPGHTAKQARRI 1777
Db	122	IALANKETLIIAGAPVYLPLVERKMGYKLLIPADSEHSALFQCCIQGVBEG-----	GLRRI 1731
Qy	178	LTLASGPPRRPIETLTSVTPDQAVAHPKMDMRKTSVDSATMMNKGLELIEACLLFNM	237
Db	174	ILTLASGAARLDLPEBLRPTVODLKLKPNMSMGKLTIDSATLNNKGLEVIEAHYLPGL	233
Qy	238	EPDQIEVATIPQSIHSMWDYVDSVLAQMGNDPMKRTPIAHAMAPERPFGVAPLDIFB	297
Db	234	DYOHIDIVIHQSIHSLIEVDIVSLAQGLMPDMRLPLLYALSBPBRITDWBEPDLVX	293
Qy	298	VGHMDEKEDLKRFPCLRLATYEAIRKSGGIMPTVLAANEIVAFLEBVKYFDIAVITZ	357
Db	294	AGLSIRREBDDHKYPCMOIAYGAGRAGGAMPVILNAANEQVALEOKISFDLPRLIB	353
Qy	358	RSMAOKFPDDAGS--LELVLOADOAR	382
Db	354	KTCDDLVGONTASPDLETTILAADQVAR	380

Search completed: January 29, 2004, 16:21:15  
Job time : 27.8953 secs